A NEW CLASS OF BIVARIATE WEIBULL DISTRIBUTIONS TO
ACCOMMODATE THE CONCORDANCE CORRELATION
COEFFICIENT FOR LEFT-CENSORED DATA

A Dissertation in
Public Health Sciences
by
Uthumporn Domthong

© 2014 Uthumporn Domthong

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

December 2014
The dissertation of Uthumporn Domthong was reviewed and approved* by the following:

Vernon M. Chinchilli  
Distinguished Professor of Public Health Sciences and Professor of Statistics  
Dissertation Advisor, Chair of Committee

Lan Kong  
Associate Professor of Public Health Sciences

Tonya S. King  
Professor of Public Health Sciences

William B. Reeves  
Professor of Nephrology, M.D.

David S. Phelps  
Professor of Pediatrics

Arthur Berg  
Associate Professor of Public Health Sciences and Bioinformatics  
Biostatistics PhD Program Director

*Signatures are on file in the Graduate School.
Abstract

In many clinical studies, Lin’s concordance correlation coefficient (CCC) is a common tool to assess the level of agreement of a continuous response measured under two different conditions. However, the complicating feature is that the assay for measuring a specific biomarker typically cannot provide accurate numerical values below the lower limit of detection (LLD), which results in left-censored data. In addition, the CCC is based on a squared distance function, and it can be very sensitive to the effects of the outliers. In this work, we propose a new class of bivariate survival functions based on functions of univariate survival functions and univariate cumulative hazard functions. We focus on using the univariate Weibull distribution to obtain a bivariate Weibull survival function. The likelihood function can be determined via this new class of bivariate Weibull survival functions. Then, we take a parametric approach to derive the estimates of the means, variances, and covariance to construct the CCC. This new class of bivariate survival functions can be extended to the situation with $p > 2$ random variables. The maximum likelihood method based on three distributions, (1) bivariate Weibull distributions (2) bivariate Farlie-Gumbel-Morgenstern distributions and (3) bivariate lognormal distributions, were evaluated via simulation studies. The simulation results confirmed that overall in terms of accuracy, that is small relative bias, the estimator of CCC based on FGM-Weibull works relatively well in general cases when the correlation is not too strong even with the high percentage of censoring. For a skewed underlying distribution with moderate or weaker correlation between two variables, the CCC estimated by a FGM-Weibull model is more robust. However, when the data are generated from the bivariate lognormal, the ML approach based on the bivariate lognormality assumption still performs best. Finally, we use data from an ancillary study of the Assessment, Serial Evaluation, and Subsequent Sequelae of Acute Kidney Injury (ASSESS-AKI) Consortium, and an asthma clinical trial for demonstration.
# Table of Contents

List of Figures vii

List of Tables viii

Acknowledgments x

Chapter 1
Introduction 1

Chapter 2
Literature Review 5
  2.1 Lin’s Concordance Correlation Coefficient 5
  2.2 CCC in the Presence of the Left-censored Data 7
  2.3 The Weibull Distribution 8
    2.3.1 Definition and Properties 8
    2.3.2 General Classes of Weibull Distributions 10

Chapter 3
Assessing the Agreement of Biomarker Data in the Presence of Left-censoring 21
  3.1 Background 21
  3.2 Methods 22
  3.3 Results 24
    3.3.1 Simulation 24
    3.3.2 Example 32
  3.4 Discussion 35
  3.5 Conclusions 36

Chapter 4
New Class of Weibull Distributions 37
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Bivariate Survival Functions</td>
<td>37</td>
</tr>
<tr>
<td>4.1.1 Bivariate Weibull-Gamma Distributions</td>
<td>43</td>
</tr>
<tr>
<td>4.1.2 Bivariate Farlie-Gumbel-Morgenstern-Weibull Distributions</td>
<td>45</td>
</tr>
<tr>
<td>4.1.3 Bivariate Piecewise Uniform-Weibull Distribution</td>
<td>49</td>
</tr>
<tr>
<td>4.2 Multivariate Survival Functions</td>
<td>51</td>
</tr>
<tr>
<td>4.2.1 Multivariate Weibull Density Function</td>
<td>52</td>
</tr>
<tr>
<td>4.2.2 Multivariate Farlie-Gumbel-Morgenstern Density Function</td>
<td>55</td>
</tr>
</tbody>
</table>

**Chapter 5**

**Simulation Studies**

5.1 Data Set Generated by a Bivariate Lognormal Distribution          61
5.2 Data Set Generated by a FGM-Weibull Distribution                  79

**Chapter 6**

**Examples**

6.1 Urine Stability Studies for Novel Biomarkers of Acute Kidney Injury 91
6.2 Asthma Clinical Data                                               101

**Chapter 7**

**Future Studies**

103

**Bibliography**

105
List of Figures

6.1 Scatter Plots of All biomarkers in Process A . . . . . . . . . . . . . . 94
6.2 Scatter Plots of All biomarkers in Process B . . . . . . . . . . . . . 95
6.3 Scatter Plots of All biomarkers in Process C . . . . . . . . . . . . . 96
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>General Classes of Weibull Distribution</td>
<td>15</td>
</tr>
<tr>
<td>2.2</td>
<td>General Classes of Weibull Distribution</td>
<td>16</td>
</tr>
<tr>
<td>2.3</td>
<td>Bivariate Weibull Distribution</td>
<td>17</td>
</tr>
<tr>
<td>2.4</td>
<td>Bivariate Weibull Distribution</td>
<td>18</td>
</tr>
<tr>
<td>2.5</td>
<td>Multivariate Weibull Distribution</td>
<td>19</td>
</tr>
<tr>
<td>2.6</td>
<td>Multivariate Weibull Distribution</td>
<td>20</td>
</tr>
<tr>
<td>3.1</td>
<td>Simulation Results Based on 1000 Data Sets with Sample Size of 100 – Percent Censoring</td>
<td>25</td>
</tr>
<tr>
<td>3.2</td>
<td>Simulation Results Based on 1000 Data Sets with Sample Size of 100 – Percent Censoring</td>
<td>26</td>
</tr>
<tr>
<td>3.3</td>
<td>Simulation Results Based on 1000 Data Sets with Sample Size of 50 – Percent Censoring</td>
<td>27</td>
</tr>
<tr>
<td>3.4</td>
<td>Simulation Results Based on 1000 Data Sets with Sample Size of 50 – Percent Censoring</td>
<td>28</td>
</tr>
<tr>
<td>3.5</td>
<td>Simulation Results Based on 1000 Data Sets with Sample Size of 25 – Percent Censoring</td>
<td>29</td>
</tr>
<tr>
<td>3.6</td>
<td>Simulation Results Based on 1000 Data Sets with Sample Size of 25 – Percent Censoring</td>
<td>30</td>
</tr>
<tr>
<td>3.7</td>
<td>Concordance Correlation Coefficients (and 95% Confidence Intervals) for 3 Processes Using 5 Methods Based on IL-18 Assay</td>
<td>33</td>
</tr>
<tr>
<td>3.8</td>
<td>Concordance Correlation Coefficients (and 95% Confidence Intervals) for 3 Processes Using 5 Methods Based on Cystatin-C Assay</td>
<td>34</td>
</tr>
<tr>
<td>5.1</td>
<td>Simulation Results Based on 1000 Data Sets with Sample Size of 100 – Percent Censoring</td>
<td>70</td>
</tr>
<tr>
<td>5.2</td>
<td>Simulation Results Based on 1000 Data Sets with Sample Size of 100 – Percent Censoring</td>
<td>71</td>
</tr>
<tr>
<td>5.3</td>
<td>Simulation Results Based on 1000 Data Sets with Sample Size of 100 – Percent Censoring</td>
<td>72</td>
</tr>
</tbody>
</table>
5.4 Simulation Results Based on 1000 Data Sets with Sample Size of 50 – Percent Censoring (25%, 25%) ....................................................... 73
5.5 Simulation Results Based on 1000 Data Sets with Sample Size of 50 – Percent Censoring (40%, 25%) ....................................................... 74
5.6 Simulation Results Based on 1000 Data Sets with Sample Size of 50 – Percent Censoring (60%, 25%) ....................................................... 75
5.7 Simulation Results Based on 1000 Data Sets with Sample Size of 25 – Percent Censoring (25%, 25%) ....................................................... 76
5.8 Simulation Results Based on 1000 Data Sets with Sample Size of 25 – Percent Censoring (40%, 25%) ....................................................... 77
5.9 Simulation Results Based on 1000 Data Sets with Sample Size of 25 – Percent Censoring (60%, 25%) ....................................................... 78
5.10 Simulation Results Based on 1000 Data Sets with Sample Size of 100 – Percent Censoring (25%, 25%) ....................................................... 82
5.11 Simulation Results Based on 1000 Data Sets with Sample Size of 100 – Percent Censoring (40%, 25%) ....................................................... 83
5.12 Simulation Results Based on 1000 Data Sets with Sample Size of 100 – Percent Censoring (60%, 25%) ....................................................... 84
5.13 Simulation Results Based on 1000 Data Sets with Sample Size of 50 – Percent Censoring (25%, 25%) ....................................................... 85
5.14 Simulation Results Based on 1000 Data Sets with Sample Size of 50 – Percent Censoring (40%, 25%) ....................................................... 86
5.15 Simulation Results Based on 1000 Data Sets with Sample Size of 50 – Percent Censoring (60%, 25%) ....................................................... 87
5.16 Simulation Results Based on 1000 Data Sets with Sample Size of 25 – Percent Censoring (25%, 25%) ....................................................... 88
5.17 Simulation Results Based on 1000 Data Sets with Sample Size of 25 – Percent Censoring (40%, 25%) ....................................................... 89
5.18 Simulation Results Based on 1000 Data Sets with Sample Size of 25 – Percent Censoring (60%, 25%) ....................................................... 90

6.1 Concordance Correlation Coefficients (and 95% Confidence Intervals) for the Three Processes Using the Bivariate Lognormal Model and Accounting for Values Below the LLD ........................................... 97
6.2 Concordance Correlation Coefficients (and 95% Confidence Intervals) for the Three Processes Using the Bivariate Weibull-Gamma Model and Accounting for Values Below the LLD ........................................... 98
6.3 Concordance Correlation Coefficients (and 95% Confidence Intervals) for the Three Processes Using the Bivariate FGM-Weibull Model and Accounting for Values Below the LLD ........................................... 99
6.4 Concordance Correlation Coefficients (and 95% Confidence Intervals) for the Three Processes Using the Bivariate Piecewise Uniform-Weibull Model and Accounting for Values Below the LLD . . . . . . . . . . . . . 100
6.5 The Two 14-week Randomized Treatment Periods . . . . . . . . . . . . 102
6.6 Frequency Counts and Cross Tabulation Table for ICS and LTRA . . 102
Acknowledgments

I would like to deeply thank my dissertation advisor Professor Vernon M. Chinchilli for his great help throughout the years that I have been working on my research. Without his guidance, support, and encouragement, this dissertation could not have been completed. I would also like to thank my other committee members, Professor Lan Kong, Professor Tonya S. King, Professor William B. Reeves, and Professor David S. Phelps for their valuable questions and comments.

In addition, I would like to thank all of those professors, staff, and colleagues in the Department of Statistics and Department of Public Health Sciences at Penn State for being so supportive during the years that I have been here. Thanks to all of my teachers who taught me well during my study in Thailand. Also thanks to all friends for the help and support.

I would like to specially thank to my family, the project DPST(2001-2006) conducted jointly by the Royal Thai Government Agencies and the Institute for the Promotion of Teaching Science and Technology(IPST), Ministry of Science and Technology (2007-2013), and the Department of Public Health Sciences (2013-2014) for their financial support. I would also like to thank Professor Vernon M. Chinchilli for giving me a research assistantship in the Public Health Sciences Department at the Penn State College of Medicine.

Last but not least, I would like to thank my mother, Mrs. Napaporn Domthong, and my siblings, Mr. Pornanan Domthong and Miss Aornchuma Domthong, for their unconditional love and care. Finally, thanks to my fiance, Mr. Thap Panitanarak, for his supports.
Dedication

To the loving memory of my grandmother, Mrs. R-Eng Sae-Tung, and my father, Mr. Anan Domthong, who are responsible for who I am and what I am today.
Biomarkers are being discovered at an accelerated rate due to availability of genomic and proteomic technologies [1]. Several of these candidate biomarkers are undergoing validation to diagnose diseases and serve as indices for predicting health outcomes. Typically, the main research interest involves comparing the biomarker result of the new technique with those of the gold standard practice. For example, Parikh et al. [2] studied on the agreement between the measurements of the urinary biomarkers between the reference standard and values obtained in samples prepared under three different processes, labeled as Process A, Process B, and Process C. In this study, the concordance correlation coefficient was invoked as the index of agreement.

The concordance correlation coefficient (CCC) was first introduced by Lin [3], and it is the most popular index of agreement. The CCC evaluates the agreement between two readings from the same sample by measuring how far each paired data point deviates from the 45 degree line through the origin, called the concordance line. To characterize the index of agreement between two random variables, $X$ and $Y$, Lin [3] considered the expected value of the squared difference, $E(X - Y)^2$, and assumed that $(X_i, Y_i), i = 1, \ldots, n$ are random sample pairs from a bivariate distribution with mean $(\mu_X, \mu_Y)$ and covariance matrix

$$
\begin{pmatrix}
\sigma_X^2 & \sigma_{XY} \\
\sigma_{XY} & \sigma_Y^2
\end{pmatrix}
$$
Lin [3] defined the CCC for $X$ and $Y$ as

$$
\rho_c = 1 - \frac{E[(X - Y)^2]}{E[(X - Y)^2 | X, Y \text{ are independent}]}
= \frac{2\sigma_{XY}}{\sigma_X^2 + \sigma_Y^2 + (\mu_X - \mu_Y)^2}
= \rho C_b
$$

where

\(\rho\) = Pearson correlation coefficient

\(C_b = \frac{2}{\nu + \frac{1}{\nu} + u^2}\)

\(\nu = \frac{\sigma_X}{\sigma_Y}\) = scale shift

\(u = \frac{(\mu_X - \mu_Y)}{\sqrt{\sigma_X \sigma_Y}}\) = location shift relative to the scale

The CCC has a scale ranging between -1 (perfect negative agreement) and +1 (perfect agreement). Zero reflects no agreement. The CCC is a more appropriate statistic than the Pearson correlation coefficient for assessing the level of agreement between two measurements of the same item. This is because the latter only quantifies the linear relationship whereas the former quantifies the linear relationship under the assumption that the slope equals one and the intercept equals zero [2].

However, when we consider bivariate data, one or both components of the paired data could be subject to censoring. For example, in practice, assays often have lower limits of detection (LLD) due to the limitation of analytic procedures, thereby making the comparison of paired values challenging. A data point below the detection limit is equivalent to being left-censored because the exact value of the data point is unknown - it only is known that it lies below the LLD. Although left-censored data are more informative than missing data, they still lead to challenges in the data analysis.

Many researchers have stressed the importance of data that are below the LLD [2, 4–6]. Hornung and Reed [6] proposed three methods of estimation with a left-censored lognormal distribution: a maximum likelihood (ML) method and two methods involving the limit of detection. However, they conclude that the ML method is too complex to calculate, so they recommend using the one-half of the LLD to represent the data that are below the LLD. Lyles et al. [4] evaluated the
Pearson’s correlation coefficient when a subset of data points was below the LLD by using the ML approach under the assumption of bivariate normality. They showed that the ML method was the most accurate among the proposed methods. Barnhart et al. [5] presented a generalized estimating equations (GEE) approach for estimating parameters to calculate the concordance correlation coefficient (CCC). The GEE approach works well and does not require the bivariate normality assumption if the sample size is large enough, and it is comparable to the ML approach when the bivariate normality assumption is appropriate. Parikh et al. [2] performed a prospective study on hospitalized patients with almost 60% of patients having acute kidney injury (AKI). Five urine biomarkers were used to compare the stability of short-term storage and processing by using the CCC as a measure of agreement. To estimate the CCC, the authors applied the ML method using log-transformed data and accounting for values below the LLD.

According to Domthong et al. [7], there are numerous methods that we can use when we encounter left-censored data. For example, simple approaches to address left-censored data are to delete value below the LLD or impute a fixed value such as one-half of the LLD or the LLD itself. However, these approaches yield biased estimates of the parameters of interest and they underestimate the variability in the data set because the same value is imputed repeatedly [4–6]. Domthong et al. [7] conclude that the ML approach with a bivariate lognormal distribution is very accurate in that it yields small relative biases, and it is best for comparing the paired data with a bivariate lognormal distribution. Therefore, in this work, we will focus on using the ML approach with bivariate data.

The ML approach is performed by constructing a likelihood function based on the bivariate distribution of the data in the detectable range, and then extrapolating into the region below the LLD. The ML approach under the assumption of bivariate normality is often used to analyze bivariate left-censored data, see Parikh et al. [2], Lyles et al. [4], and Hornung and Reed [6]. In this work, we describe the ML approach and present an alternative Weibull distribution to evaluate agreement between two assays that are both subject to lower limits of detection.

Several classes of bivariate and multivariate Weibull distributions have been proposed previously. For example, the multivariate Weibull distribution generated from correlated Gaussian processes was first introduced by Sagias and Karagiannidis [8]. Franco and Vivo [9] proposed a generalized mixture of Weibull distribution with
common shape parameter. A Marshall-Olkin bivariate Weibull (MOBW) distribution was summarized in Nandi and Dewan [10]. The Farlie-Gumbel-Morgenstern copula is the one of the most popular parametric families of copulas that was studied in Farlie [11], Gumbel [12], and Morgenstern [13]. Hougaard [14] proposed a class of continuous multivariate lifetime distributions with a survival function. The latter class is a copula that has been studied in many research areas. For example, See Gumbel, Clayton, Frank and Independence copulas in Lee et al. [15], Gaussian copula in Song et al. [16], and Farlie-Gumbel-Morgenstern copulas in Bekrizadeh et al. [17]. However, no class of multivariate Weibull distributions generated from the survival function and the cumulative hazard function has ever been published. In this work, we propose a new class of multivariate Weibull distributions, and use them to apply the maximum-likelihood based approach.

In chapter 2, we provide more detail on Lin’s CCC. The previous work on CCC in the presence of the left-censored data are reviewed as well. In addition, definitions and properties of Weibull distribution are summarized, along with several classes of bivariate and multivariate Weibull distributions. In chapter 3, the literature on the assessing the agreement of biomarker data in the presence of left-censoring is reviewed. In chapter 4, we state the expression for the joint density and survival function of the new class of multivariate Weibull distributions. In chapter 5, simulation studies are performed to check flexibility, accuracy, and relative robustness of the procedures. Finally, we use data from an ancillary study of the Assessment, Serial Evaluation, and Subsequent Sequelae of Acute Kidney Injury (ASSESS-AKI) Consortium and an Asthma Clinical trial for a demonstration in chapter 6, and the direction of future work is discussed in chapter 7.
Chapter 2 | Literature Review

In this chapter, the concordance correlation coefficient which was introduced by Lin [3], and the previous works of the CCC with left-censored data are reviewed. Then, the Weibull distribution and some useful properties are reviewed. In addition, the general classes of the Weibull distribution, the advantages, and the disadvantages for each class of Weibull distributions are summarized.

2.1 Lin’s Concordance Correlation Coefficient

Assume that $(X_i, Y_i), i = 1, \ldots, n$ are random sample pairs from a bivariate distribution with mean $(\mu_X, \mu_Y)$ and covariance matrix

$$
\begin{pmatrix}
\sigma_X^2 & \sigma_{XY} \\
\sigma_{XY} & \sigma_Y^2
\end{pmatrix}
$$

Lin [3] defined the CCC for $X$ and $Y$ as

$$
\rho_c = 1 - \frac{E[(X - Y)^2]}{E[(X - Y)^2 | X, Y \text{ are independent}]}
= \frac{2\sigma_{XY}}{\sigma_X^2 + \sigma_Y^2 + (\mu_X - \mu_Y)^2}
= \rho C_b
$$

(2.1)
where
\[ \rho = \text{Pearson correlation coefficient} \]
\[ C_b = \frac{2}{\nu + \frac{1}{\nu} + u^2} \]
\[ \nu = \frac{\sigma_X}{\sigma_Y} = \text{scale shift} \]
\[ u = \frac{(\mu_X - \mu_Y)}{\sqrt{\sigma_X\sigma_Y}} = \text{location shift relative to the scale} \]

The Pearson correlation coefficient measure how far the observations deviate from the fitted line (measures of precision), and \(-1 \leq \rho \leq +1\). On the other hand, \(C_b\) is a bias correction factor that measures how far the best-fit line deviates from the concordance line (measure of accuracy), and \(0 < C_b \leq 1\). The further \(C_b\) is from 1, the greater the deviation is from the concordance line. When \(C_b = 1\), no deviation occurs.

The concordance correlation coefficient evaluates the degree to which pairs fall on the concordance line, and \(-1 \leq -|\rho| \leq \rho_c \leq |\rho| \leq 1\). Any departure from this line would produce \(\rho_c < 1\) even if \(\rho = 1\). Moreover, \(\rho_c = 0\) iff \(\rho = 0\), and \(\rho_c = \pm 1\) iff \(\rho = \pm 1, \sigma_X = \sigma_Y\), and \(\mu_X = \mu_Y\), or equivalently, \(\rho_c = \pm 1\) iff each observations falls on the concordance line.

Lin [3] used the sample counterparts to estimate \(\rho_c\) as follows,
\[ \hat{\rho}_c = \frac{2S_{XY}}{S_X^2 + S_Y^2 + (\bar{X} - \bar{Y})^2} \quad (2.2) \]

where
\[ \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i, \quad \bar{Y} = \frac{1}{n} \sum_{i=1}^{n} Y_i, \]
\[ S_X^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2, \quad S_Y^2 = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \bar{Y})^2 \]
\[ S_{XY} = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y}) \]

Note that Lin [3] didn’t use the unbiased estimators for the sample counterparts to estimate \(\rho_c\), but we will use the unbiased estimators that are comparable to Lin’s,
that is, the sample counterparts that we use in this work are as follow:

\[ S^2_X = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2, \quad S^2_Y = \frac{1}{n-1} \sum_{i=1}^{n} (Y_i - \bar{Y})^2 \]

\[ S_{XY} = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y}) \]

Assume that \( \hat{\rho}_c \) is the sample concordance correlation coefficient of paired samples from a bivariate normal distribution. Firstly, by using the Z-transformation, Lin showed that

\[ \hat{Z} = \tanh^{-1}(\hat{\rho}_c) = \frac{1}{2} \ln \frac{1 + \hat{\rho}_c}{1 - \hat{\rho}_c} \]

is asymptotic normal with mean \( Z = \frac{1}{2} \ln \frac{1 + \rho_c}{1 - \rho_c} \) and variance

\[ \sigma_Z^2 = \frac{1}{n-2} \left[ \frac{(1 - \rho^2)\rho_c^2}{(1 - \rho_c^2)\rho^2} + \frac{4\rho_c^3(1 - \rho_c)u^2}{\rho(1 - \rho_c^2)^2} - \frac{2\rho_c^4u^4}{\rho^2(1 - \rho_c^2)^2} \right] \]

Then, by using the above transformation, it can be shown that \( \hat{\rho}_c \) is a consistent estimator of \( \rho_c \) and has an asymptotic normal distribution with with mean \( \rho_c \) and variance \( (1 - \rho_c^2)^2 \sigma_Z^2 \).

### 2.2 CCC in the Presence of the Left-censored Data

Many researchers have stressed the importance of data that are below the LLD [2,4–6]. First, Hornung and Reed [6] proposed three methods of estimation with a left-censored lognormal distribution: a maximum likelihood (ML) method and two simple data imputation methods. However, they conclude that the ML method is too complex to calculate, so they recommend using one-half of the LLD to represent the left-censored data.

Lyles et al. [4] assessed the Pearson’s correlation coefficient between two variables, both of which are subject to left censoring due to some values falling below assay detection limits by using a maximum likelihood (ML) approach based on the bivariate normality assumption.

To extend the previous work, Barnhart et al. [5] used the ML approach under a normality assumption to estimate the concordance correlation coefficient (CCC) and
also presented the generalized estimating equations (GEE) approach for estimating the CCC for the left-censored data. They concluded that the GEE approach works well and comparable to the ML approach based on the bivariate normality assumption.

Parikh et al. [2] proposed another version of maximum likelihood based approach by using a bivariate lognormal distribution. This approach was illustrated by using a subset of the data from the Assessment, Serial Evaluation, and Subsequent Sequelae in Acute Kidney Injury Study (ASSESS-AKI). The objective of the statistical analysis was to determine the agreement between the measurements of the urinary biomarkers between the reference standard and values obtained in samples prepared under different conditions (labeled as Process A, Process B, and Process C). Lin’s CCC was used to assess the level of agreement of biomarkers under two different conditions. Instead of using Pearson’s correlation coefficient, CCC was used in this study because it quantifies the linear relationship under the assumption that the slope equals one and the intercept equals zero. On the other hand, Pearson’s correlation coefficient only quantifies the linear relationship. A likelihood function was constructed based on the bivariate lognormal distribution in which values below the LLD were included as left-censored values. This reduced the number of samples deleted from analysis. Our study parallels an earlier report in Parikh et al. [2].

2.3 The Weibull Distribution

2.3.1 Definition and Properties

Johnson and Kotz [18] defined that a random variable $X$ has a Weibull distribution if there are values of the parameter $\beta(>0), \lambda(>0)$ and $\mu_0$ such that

$$Y = [\lambda(X - \mu_0)]^\beta$$

has the exponential distribution, with the probability density function

$$f_Y(y) = \exp(-y) \quad \text{where} \quad y > 0$$

The probability density function of $X$ is

$$f_X(x) = \lambda \beta [\lambda(x - \mu_0)]^{\beta-1} \exp(-[\lambda(x - \mu_0)]^\beta) \quad \text{where} \quad \mu_0 < x \quad (2.3)$$
The standard Weibull distribution is obtained by putting \( \mu_0 = 0 \) and \( \lambda = 1 \).

In the remainder of this work, all the distributions to be discussed will have \( \mu_0 = 0 \) and \( \lambda > 0 \). Also, we assume \( x > 0 \) and \( \beta > 0 \). Therefore, we have that the probability density function of a univariate Weibull random variable \( X \) is

\[
f_X(x) = \lambda \beta x^{\beta-1} \exp(-\lambda x^\beta)
\]  

(2.4)

where \( x > 0, \lambda > 0, \) and \( \beta > 0 \).

Note that (2.3) and (2.4) are the same form. We can simply see this re-parameterization by replacing \( \lambda \) in (2.3) with \( \lambda^\beta \). However, the form (2.4) will be used throughout this work.

The cumulative distribution function for the univariate Weibull distribution is

\[
F_X(x) = 1 - \exp(-\lambda x^\beta)
\]

Therefore, the survival function for the univariate Weibull distribution is

\[
S_X(x) = \exp(-\lambda x^\beta)
\]

The failure rate \( h \) (or hazard rate) is given by

\[
h_X(x) = \frac{f_X(x)}{S_X(x)} = \lambda \beta x^{\beta-1}
\]

Properties

According to [18], the expected value and variance of a Weibull random variable \( X \) in (2.3) can be expressed as

\[
E(X) = \left( \frac{1}{\lambda} \right)^{\frac{1}{\beta}} \Gamma \left( 1 + \frac{1}{\beta} \right)
\]

and

\[
Var(X) = \left( \frac{1}{\lambda} \right)^{\frac{2}{\beta}} \left\{ \Gamma \left( 1 + \frac{2}{\beta} \right) - \Gamma^2 \left( 1 + \frac{1}{\beta} \right) \right\}
\]

where

\[
\Gamma(\alpha) = \int_0^\infty \exp(-x)x^{\alpha-1} \, dx, \quad \alpha > 0
\]
2.3.2 General Classes of Weibull Distributions

A Weibull distribution is well-known and widely used in many research areas. There are numerous extensions of Weibull distributions to fit data and model. There are six major classes of Weibull distributions that we summarize: (1) Gaussian Class Multivariate Weibull Distribution, (2) Mixture of Weibull Distribution, (3) Marshall-Olkin Weibull, (4) Farlie-Gumbel-Morgenstern, (5) Hougaard, and (6) Copula. In this dissertation, we summarize all general classes of Weibull distributions, the advantages, and the disadvantages in Table 2.1 and Table 2.2. Last but not least, we also summarize the bivariate and multivariate Weibull distributions for all classes in Table 2.3, Table 2.4 and Table 2.5.

1. Gaussian Class Multivariate Weibull Distributions

The multivariate Weibull distribution generated from correlated Gaussian processes was first introduced by Sagias and Karagiannidis [8] for analyzing the performance of diversity receivers operating over more realistic correlated fading channels. Specifically, the bivariate Weibull pdf with not necessarily identical fading parameters as well as average powers is presented. While based on this pdf, the corresponding moments-generating function (mgf), cdf, and the Weibull correlation coefficient are obtained. Multivariate Weibull distributions with exponential and constant correlation matrices are also introduced and for the former, useful analytical expressions for the joint pdf, cdf, mgf, and product moments are presented. These novel theoretical results are applied to the performance analysis of dual- and multibranch SC, EGC, and MRC receivers operating over correlated Weibull fading channels as appeared in [8].

The bivariate probability density function of Gaussian class was derived by Sagias and Karagiannidis [8] as:

\[
f(x_1, x_2) = \frac{\beta_1 \beta_2 x_1^{\beta_1 - 1} x_2^{\beta_2 - 1}}{\Omega_1 \Omega_2 (1 - \rho)} \times \exp \left[ - \frac{1}{1 - \rho} \left( \frac{x_1^{\beta_1}}{\Omega_1} + \frac{x_2^{\beta_2}}{\Omega_2} \right) \right] \times I_0 \left( \frac{2 \sqrt{\rho x_1^{\beta_1/2} x_2^{\beta_2/2}}}{(1 - \rho) \sqrt{\Omega_1 \Omega_2}} \right)
\]

where \( \Omega_l = E(X_l^{\beta_l}) \) for \( l = 1 \) and 2.

Let \( X = [X_1 X_2 \ldots X_p]^T \). Then the multivariate probability density function of
Gaussian class of Weibull distributions is

\[
f(X) = \frac{1}{\Omega_P} \prod_{i=1}^{p} \beta_i x_i^{\beta_i - 1} \times \exp \left( - \frac{1}{(1 - \rho)^\Omega} \left[ x_1^{\beta_1} + x_p^{\beta_p} + (1 + \rho) \sum_{i=2}^{p-1} x_i^{\beta_i} \right] \right) \\
\times \prod_{i-1}^{p-1} I_0 \left[ \frac{2\sqrt{\rho}}{(1 - \rho)^\rho} x_i^{\beta_i/2} x_{i+1}^{\beta_{i+1}/2} \right]
\]

2. Mixture of Weibull Distributions

Let \((X_1, X_2, \ldots, X_n)\) be a random vector formed by Weibull components with common shape parameter \(c > 0\) and density function \(f_i(x) = c b_i x^{c-1} \exp(-b_i x^c)\) for all \(x > 0\), and \(b_i > 0, i = 1, 2, \ldots, n\). Then Franco and Vivo [9] define that \(X\) is a generalized mixture of Weibull distributions with common shape parameter \(c > 0\) if its density function (pdf) is given by

\[
f(x) = \sum_{i=1}^{n} a_i f_i(x)
\]

where \(a_i \in \mathbb{R}, i = 1, 2, \ldots, n\), such that \(\sum_{i=1}^{n} a_i = 1\).

This generalized mixture may be defined by its distribution and survival functions, respectively,

\[
F(x) = 1 - \sum_{i=1}^{n} a_i \exp(-b_i x^c) \quad \text{and} \quad S(x) = \sum_{i=1}^{n} a_i \exp(-b_i x^c)
\]

A brief review of the generalized mixtures of exponential distributions and properties related with these generalized mixtures can be found in [9]. The interesting results investigated in [9] were constraints on the mixing weights and parameters of components under which the generalized mixture of Weibull distributions is a valid probability model, including the cases of exponential or Rayleigh components. In addition, characterizations are shown for generalized mixtures of three or fewer Weibull components.

3. Marshall-Olkin Weibull

Marshall and Olkin [19] first introduced the well-known multivariate
exponential distribution (MVE) with survival function

$$
\Pr(X_1 > x_1, \ldots, X_n > x_n) = \exp \left\{ - \sum_{i=1}^{n} \lambda_i x_i - \sum_{i<j} \lambda_{i,j} \max(x_i, x_j) \\
- \sum_{i<j<k} \lambda_{i,j,k} \max(x_i, x_j, x_k) - \cdots - \lambda_{1,\ldots,n} \max(x_i : i \in I) \right\}
$$

for $x_1, \ldots, x_n \geq 0$ and $\lambda_I \geq 0$ for any $I \subseteq \{1, \ldots, n\}$.

Afterward, insightful properties and results were proven to get a better understanding of this MVE distributions, see Galambos and Kotz [20] and Muliere and Scarsini [21]. For some recent work, there are many interesting Marshall-Olkin Bivariate Weibull (MOBW) distributions, and we refer readers to Nandi and Dewan [10]. They derive the expressions for the joint density and survival function of MOBW distribution for completeness, and study the likelihood function of data coming from MOBW distribution when it is subjected to random right censoring. Li and Pellerey [22] introduced the bivariate generalized Marshall-Olkin distribution and had a study on the evolution of the dependence with respect to the age and the related aging properties as well. The extension of the main results in Li and Pellerey [22] was derived by Lin and Li [23]. They derived high dimensional generalizations of some results on the bivariate generalized Marshall-Olkin distributions.

4. Farlie-Gumbel-Morgenstern (FGM)

One of the most popular parametric families of copulas studied in Farlie [11], Gumbel [12], and Morgenstern [13], is the Farlie-Gumbel-Morgenstern (FGM) family defined by

$$
F_{X,Y}(x, y) = F_X(x)F_Y(y)[1 + \alpha(1 - F_X(x))(1 - F_Y(y))] \quad \text{where} \quad |\alpha| \leq 1 \quad (2.6)
$$

and the FGM copula density is provided by

$$
f_{X,Y}(x, y) = f_X(x)f_Y(y)[1 + \alpha(2F_X(x) - 1)(2F_Y(y) - 1)] \quad \text{where} \quad |\alpha| \leq 1 \quad (2.7)
$$

where $F_X$, $F_Y$, $F_X$, and $f_Y$ are the marginal cumulative distribution functions and marginal density functions, respectively, and $F_{XY}$ and $f_{XY}$ are the bivariate cumulative distribution function and bivariate density function, respectively.
A bivariate copula can be statistically interpreted as a bivariate distribution function with uniform marginals. Let $X$ and $Y$ be random variables where $X \sim \text{Weibull}(\lambda_1, \beta_1)$ where $\lambda_1 > 0, \beta_1 > 0$. Then $F_X(x) = 1 - \exp(-\lambda_1 x^{\beta_1})$

$Y \sim \text{Weibull}(\lambda_2, \beta_2)$ where $\lambda_2 > 0, \beta_2 > 0$. Then $F_Y(y) = 1 - \exp(-\lambda_2 y^{\beta_2})$

Then there exists a Farlie-Gumbel-Morgenstern copula such that the bivariate distribution function $F_{X,Y}(x,y)$ is the following:

$$F_{X,Y}(x,y) = F_X(x)F_Y(y)[1 + \alpha(1 - F_X(x))(1 - F_Y(y))] \quad \text{where } |\alpha| \leq 1$$

$$= \left[1 - \exp(-\lambda_1 x^{\beta_1})\right] \times \left[1 - \exp(-\lambda_2 y^{\beta_2})\right] \times \left[1 + \alpha \exp(-\lambda_1 x^{\beta_1} - \lambda_2 y^{\beta_2})\right]$$

Moreover, we can easily see that the marginal distributions of this bivariate distribution are Weibull distributions and are independent when $\alpha = 0$.

5. **Hougaard**

Hougaard [14] proposed a class of continuous multivariate lifetime distributions with a survival function

$$S(x_1, x_2, \ldots, x_k) = \exp\left(-\left(\sum_{i=1}^{k} \lambda_i x_i^{\beta_i}\right)^\gamma\right).$$

A substantial amount of literature on various bivariate and multivariate exponential models is given in Basu [24]. More recently, Mohsin et al. [25] introduced a new bivariate exponential distribution. The properties of this distribution were presented along with estimation procedures for the model parameters based on maximum likelihood and objective Bayesian analysis.

6. **Copulas**

According to Lee et al. [15], copulas are a useful tool for developing a joint distribution function. In particular, copulas have gained their importance as simple functions to describe the dependence structure of random variables in the joint distribution. In their paper, they also described several advantages of copulas over other dependence measures, and the use of copulas in applications. For example, using copulas, modeling both linear and non-linear dependencies of
variables is possible, and the degree of dependence in the tail of the underlying distribution can be described as mentioned in Embrechts et al. [26]. Copula was applied to use in risk management and survival analysis research area. There are a lot of families of copulas that have been studied in many research areas. Example are Gumbel, Clayton, Frank and Independence copulas in Lee et al. [15], Gaussian copula in Song et al. [16], Farlie-Gumbel-Morgenstern copulas in Bekrizadeh et al. [17].

According to Lee et al. [15], given two marginal Weibull distributions $F_i(x_i) = 1 - \exp(-\lambda_i x_i^{\beta_i}), i = 1, 2$, it is possible to construct a bivariate distribution:

- Gumbel copula $F(x_1, x_2) = \exp \left\{ -\left[ (-\log F_1(x_1))^\theta + (-\log F_2(x_2))^\theta \right]^{1/\theta} \right\}$
- Clayton copula $F(x_1, x_2) = \left[ F_1(x_1)^{-\theta} + F_2(x_2)^{-\theta} - 1 \right]^{-1/\theta}$
- Frank copula $F(x_1, x_2) = -\frac{1}{\theta} \log \left[ 1 + \frac{(e^{-\theta F_1(x_1)} - 1)(e^{-\theta F_2(x_2)} - 1)}{e^{-\theta} - 1} \right]$
- Independence copula $F(x_1, x_2) = F_1(x_1) F_2(x_2)$
<table>
<thead>
<tr>
<th>Classes</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Gaussian Class Multivariate Weibull Distribution</td>
<td>The bivariate Weibull pdf with not necessarily identical fading parameters as well as average powers is presented, introduced by Sagias and Karagiannidis [8]</td>
<td>The marginal distributions are not necessarily Weibull distributions.</td>
</tr>
<tr>
<td>2. Mixture of Weibull Distribution</td>
<td>Franco and Vivo [9] mentioned that Weibull mixture models play an important role in reliability theory, since they exhibit a wide range of shapes for the density and failure rate functions, which makes them suitable for modeling complex survival or failure data sets, and they can be generalized by allowing negative mixing weights.</td>
<td>There are complex restrictions on the mixing weights and parameters of components under which the generalized mixture of Weibull distributions is a valid probability model.</td>
</tr>
<tr>
<td>3. Marshall-Olkin Weibull</td>
<td>According to Nandi and Dewan [10], this distribution fits a bivariate data set very well if it has a unimodal marginal density function or has a non-constant hazard function. Besides, it is often used to fit paired data in survival studies where there is a possibility of simultaneous occurrence of both the events.</td>
<td>As mentioned in Lin and Li [23], the essential difference of the dependence between two random vectors in the sense of some stochastic orders for the survival copula of a multivariate generalized Marshall-Olkin distribution is still difficult to build a comparison result through two explicit expressions.</td>
</tr>
<tr>
<td>4. Farlie-Gumbel-Morgenstern</td>
<td>This class is a special case of a copula. Therefore, the property is the same as a copula.</td>
<td>By using the admissible range of $\theta$, a well-known limitation to FGM copula, is that it does not allow the modeling of high dependences, since Spearman’s correlation coefficient is limited to $\rho \in (-0.33, 0.33)$. But by this generalization, Bekrizadeh et al. [17] show the range of Spearman’s $\rho$ can be wider than with respect to Spearman’s $\rho$ in FGM copulas.</td>
</tr>
</tbody>
</table>
Table 2.2. General Classes of Weibull Distribution

<table>
<thead>
<tr>
<th>Classes</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>5. ( S(x_1, x_2) = \exp(- (\lambda_1 x_1^\beta + \lambda_2 x_2^\beta)\gamma) )</td>
<td>Hougaard [14] showed that this model has closure properties similar to those of the normal variance components model. The result of Lemma in this paper implies that ( \text{Corr}(\log X_1, \log X_2) = 1 - \gamma^2 ), independently of ( \lambda_1, \lambda_2, ) and ( \beta )</td>
<td>This model is not be rich enough to extend to the multivariate form because it is not clear how to be generalized and also only 1 parameter used to describe the relationship among ( p ) variables when ( p \geq 2 ).</td>
</tr>
<tr>
<td>6. Copula</td>
<td>Using copulas, modeling both linear and non-linear dependencies of variables is possible. In addition, the degree of dependence in the tail of the underlying distribution can be described, mentioned by Lee et al. [15].</td>
<td>According to [15], copulas have varying amounts of tail dependence depending on the choice of the copula. Therefore, an important issue in using copulas is the choice of an appropriate copula. Poorly chosen copulas may lead to undesired results about the actual relationship between variables.</td>
</tr>
</tbody>
</table>
Table 2.3. Bivariate Weibull Distribution

<table>
<thead>
<tr>
<th>Classes</th>
<th>Bivariate probability density function/ survival function/ cumulative distribution function of Weibull Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Gaussian Class Multivariate Weibull Distribution</td>
<td>$f(x_1, x_2) = \frac{\beta_1 \beta_2 x_1^{\beta_1-1} x_2^{\beta_2-1}}{\Omega_1 \Omega_2 (1-\rho)} \times \exp \left[-\frac{1}{1-\rho} \left(x_1^{\beta_1} + x_2^{\beta_2}\right)\right] I_0 \left[\frac{2 \sqrt{\pi} x_1^{\beta_1/2} x_2^{\beta_2/2}}{(1-\rho) \sqrt{\Omega_1 \Omega_2}}\right]$ where $\Omega_l = E(X_l^{\beta_l})$ for $l = 1$ and $2$, derived by Sagias and Karagiannidis [8].</td>
</tr>
<tr>
<td>2. Mixture of Weibull Distribution</td>
<td>$f(x) = cx^{c-1} \left(a_1 b_1 \exp(-b_1 x^c) + a_2 b_2 \exp(-b_2 x^c)\right)$ be a generalized mixture of two Weibull density functions with common shape parameter $c &gt; 0$, $0 &lt; b_1 &lt; b_2 &lt; \infty$, $a_1 &gt; 0$ and $a_2 \in \mathbb{R}$ such that $a_1 + a_2 = 1$, and either $a_2 &gt; 0$ or $a_1 b_1 + a_2 b_2 \geq 0$, shown by [9].</td>
</tr>
</tbody>
</table>
| 3. Marshall-Olkin Weibull                   | According to Nandi and Dewan [10], the density function is given by $f_{WE}(x; \alpha, \theta) = \alpha \theta x^{\alpha-1} \exp(-\theta x^\alpha)$. The joint density function of $(X_1, X_2)$ is given as:
- $f(x_1, x_2) = f_{WE}(x_1; \alpha, \lambda_1) f_{WE}(x_2; \alpha, \lambda_0 + \lambda_2)$ if $x_1 < x_2$
- $f(x_1, x_2) = f_{WE}(x_1; \alpha, \lambda_0 + \lambda_1) f_{WE}(x_2; \alpha, \lambda_2)$ if $x_1 > x_2$
- $f(x_1, x_2) = \frac{\lambda_0}{x} f_{WE}(x; \alpha, \lambda)$ if $x_1 = x_2$

4. Farlie - Gumbel - Morgenstern             | From (2.7), we have $f_{X_1, X_2}(x_1, x_2) = \left[1 + \alpha(1 - 2 \exp(-\lambda_1 x_1^{\beta_1}))(1 - 2 \exp(-\lambda_2 x_2^{\beta_2}))\right]$ where $|\alpha| \leq 1$

5. Hougaard                                  | According to Hougaard [14], we have $S(x_1, x_2) = \exp(-\lambda_1 x_1^{\beta_1} + \lambda_2 x_2^{\beta_2})$). |
Table 2.4. Bivariate Weibull Distribution

<table>
<thead>
<tr>
<th>Classes</th>
<th>Bivariate probability density function/ survival function/ cumulative distribution function of Weibull Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>6. Copula Based</td>
<td>According to Lee et al. [15], given two marginal Weibull distributions $F_i(x_i) = 1 - \exp(-\lambda_i x_i^\beta_i)$, $i = 1, 2$, it is possible to construct a bivariate distribution:</td>
</tr>
<tr>
<td></td>
<td>• Gumbel copula $F(x_1, x_2) = \exp\left{(-\log F_1(x_1))^\theta + (-\log F_2(x_2))^\theta\right}^{1/\theta}$</td>
</tr>
<tr>
<td></td>
<td>• Clayton copula $F(x_1, x_2) = (F_1(x_1)^{-\theta} + F_2(x_2)^{-\theta} - 1)^{-1/\theta}$</td>
</tr>
<tr>
<td></td>
<td>• Frank copula $F(x_1, x_2) = -\frac{1}{\theta} \log\left[1 + \frac{(e^{-\theta F_1(x_1)} - 1)(e^{-\theta F_2(x_2)} - 1)}{e^{-\theta} - 1}\right]$</td>
</tr>
<tr>
<td></td>
<td>• Independence copula $F(x_1, x_2) = F_1(x_1)F_2(x_2)$</td>
</tr>
</tbody>
</table>
Table 2.5. Multivariate Weibull Distribution

<table>
<thead>
<tr>
<th>Classes</th>
<th>Multivariate probability density function / survival function / cumulative distribution function of Weibull Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Gaussian Class Multivariate Weibull Distribution</td>
<td>Let $X = [X_1 X_2 \ldots X_p]^T$. Then $f(X) = \frac{1}{(1-\rho)^{p\beta_1-1}} \times \exp \left( -\frac{1}{(1-\rho)^{p\beta_1}} x_1^\beta_1 + x_p^\beta + (1 + \rho) \sum_{i=2}^{p-1} x_i^{\beta_i} \right) \times \prod_{i=1}^{p-1} I_0 \left[ \frac{2\sqrt{\pi}}{(1-\rho)^{p\beta_i/2}} x_i^{\beta_i/2} x_{i+1}^{\beta_i+1/2} \right]$, derived by Sagias and Karagiannidis [8]</td>
</tr>
<tr>
<td>2. Generalized mixture of Weibull Distribution</td>
<td>Let $X = [X_1 X_2 \ldots X_p]^T$, and density function $f_i(x) = cb_i x^{c-1} \exp(-b_i x^c)$ for $x &gt; 0$, and $b_i &gt; 0, i = 1, 2, \ldots, n$. Then $X$ is a generalized mixture of Weibull Distribution with common shape parameter $\beta &gt; 0$ if $f(x) = \sum_{i=1}^{n} a_i f_i(x)$ where $a_i \in \mathbb{R}, i = 1, 2, \ldots, n$, such that $\sum_{i=1}^{n} a_i = 1$, defined by Franco and Vivo [9]</td>
</tr>
<tr>
<td>3. Marshall-Olkin</td>
<td>Marshall-Olkin type survival function is $Pr(X_1 &gt; x_1, \ldots, X_n &gt; x_n) = \exp \left( \sum_{i=1}^{n} \lambda_i H(x_i) - \sum_{i&lt;j} \lambda_{i,j} H(x_i \vee x_j) - \sum_{i&lt;j&lt;k} \lambda_{i,j,k} H(x_i \vee x_j \vee x_k) - \cdots - \lambda_{i,\ldots,n} H(\vee_{i=1}^{n} x_i) \right)$ for $x_1, \ldots, x_n \geq 0$ and increasing $H$ with $H(0) = 0$ and $H(+\infty) = +\infty$ and $\lambda \geq 0$ for any $I \subseteq (1, \ldots, n)$, here $x \vee y = \max(x, y), i \in It_i = \max(t_i : i \in I)$, according to Muliere and Scarsini [21],</td>
</tr>
<tr>
<td>4. Farlie - Gumbel - Morgenstern</td>
<td>A multivariate FGM cumulative distribution function as $F_{12\ldots p}(x_1, x_2, \ldots, x_p) = F_1(x_1) F_2(x_2) \ldots F_p(x_p) \left( 1 + \alpha_{12} S_1(x_1) S_2(x_2) \right) \times \left( 1 + \alpha_{13} S_1(x_1) S_3(x_3) \right) \times \cdots \times \left( 1 + \alpha_{p-1,p} S_{p-1}(x_{p-1}) S_p(x_p) \right)$ where each $\alpha_{ij}$ is a correlation parameter, $-1 \leq \alpha_{ij} \leq +1, i = 1, 2, \ldots, p-1$ and $j = i + 1, i + 2, \ldots, p$. Where $F_i(x_i) = \int_0^{x_i} f_i(t) , dt$ denote the cumulative distribution function $S_i(x_i) = 1 - F_i(x_i) = \int_{x_i}^{\infty} f_i(t) , dt$ denote the survival function</td>
</tr>
</tbody>
</table>
Table 2.6. Multivariate Weibull Distribution

<table>
<thead>
<tr>
<th>Classes</th>
<th>Multivariate probability density function / survival function / cumulative distribution function of Weibull Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>5. Hougaard</td>
<td>According to Hougaard [14], we have $S(x_1, x_2, \ldots, x_k) = \exp\left(-\left(\sum_{i=1}^{k} \lambda_i x_i^\beta\right)^\gamma\right)$</td>
</tr>
<tr>
<td>6. Copula Based</td>
<td>A copula function, denoted by $C$, is a multivariate distribution function with uniform marginal distribution functions, $F_1, \ldots, F_p$ on the interval $[0, 1]$, i.e., if for $x_1, x_2, \ldots, x_p, F_{12\ldots p}(x_1, x_2, \ldots, x_p)$ is a multivariate probability distribution with marginals $F_1(x_1), F_2(x_2), \ldots, F_p(x_p)$ then $F_{12\ldots p}(x_1, x_2, \ldots, x_p)$ can be written as $F_{12\ldots p}(x_1, x_2, \ldots, x_p) = C(F_1(x_1), F_2(x_2), \ldots, F_p(x_p))$</td>
</tr>
</tbody>
</table>
Chapter 3  |  Assessing the Agreement of Biomarker Data in the Presence of Left-censoring

In order to determine an efficient method for analyzing left-censored data, we study the performance of several methods such as the data deletion method, simple data imputation method, and the maximum likelihood (ML) method under the assumption of a bivariate lognormal distribution. We conduct a computer simulation study to investigate the statistical properties of the ML method versus the data deletion and simple data imputation method. The study resulted in context-specific assessment of performance of the Lin’s Concordance Correlation Coefficient (CCC) [3]. The main conclusions reported in this chapter are taken from Domthong et al. [7].

3.1  Background

Biomarkers in blood and urine are important indicators for the diagnosis of diseases and risk-stratification. During the development of biomarker assays, several pre-analytic steps require comparison of paired values of biomarkers exposed to separate conditions, such as varying degrees of storage, freeze-thaw cycles, and different antibodies. For these experiments, comparison of paired biomarker values is a critical step to advance the development to the next stage. In practice, assays often have lower limits of detection (LLD) due to the limitation of analytic procedures, thereby making the comparison of paired values challenging. A data point below the
detection limit is equivalent to being left censored because the exact value of the data point is unknown - it only is known that it lies below the LLD. Although left-censored data are more informative than missing data, they still lead to challenges in the data analysis.

Simple (ad hoc) approaches to address the left-censored data are to delete the value below the LLD or impute a fixed value such as one-half of the LLD or the LLD itself. However, these approaches yield biased estimates of the parameters of interest and they underestimate the variability in the data set because the same value is imputed repeatedly [4–6]. Urine biomarkers are very prone to this problem as the concentration of the biomarkers is greatly influenced by urine volume. In diluted urine, biomarker values may be below the LLD. Also, biomarkers whose concentrations are below ng levels are prone to this problem of having values below LLD. For example, IL-18 is measured in pg/vol in urine, and thus usually has higher proportions of values below the LLD compared to other biomarkers. Many researchers have stressed the importance of data that are below the LLD [2,4–6].

From the statistical point of view, we can expect the ad hoc methods (data deletion or simple imputation) to estimate the data differently and in a biased manner from the ML approach. An ideal approach to handle the left-censored data is to invoke the ML method because it accounts for the distribution of data in the detectable range and extrapolates into the region below the LLD.

The aim of this study is to show that when faced with left-censored data, the ML approach based on a bivariate normal (or lognormal) assumption for estimating the CCC between two assays is a more appropriate approach to use in practice than the ad hoc approaches that involve data deletion or simple data imputation.

3.2 Methods

To find the optimal method to deal with left-censored data, we investigate how data deletion and simple data imputation methods compare to the ML approach in a computer simulation study. We adapt the framework from Barnhart et al. [5] for our computer simulation studies. In all simulations described in the results, we generate bivariate normal data for paired data represented by the variables $X$ and $Y$ with a sample size of 100, 50, 25 for each of 1000 data sets using one of the
following six combinations of parameter settings for the means, standard deviations, and correlation coefficient: $\mu_x = 0, \mu_y = 0.2, \sigma_x = 0.8, \sigma_y = 1, \rho = 0.25, 0.50, 0.75$, and left-censoring rates of (25% for X, 25% for Y) or (40% for X, 25% for Y). The selected values of the LLDs in the simulation study are determined by the censoring rates. All calculations are performed using SAS 9.3 statistical software. All estimated CCCs ($\hat{\rho}_c$) were obtained by maximizing the likelihood function with respect to each of the following five scenarios.

1. Deleting the pair method means that pairs with X, Y, or both X and Y below the detection limit are discarded before calculation of the CCC. The 95% confidence interval (CI) of this method is calculated by using $\hat{\rho}_c \pm Z_{0.025}SE(\hat{\rho}_c)$ where $\hat{\rho}_c$ is the estimated CCC, $Z_{0.025}$ is the critical value of the standard normal distribution, and $SE(\hat{\rho}_c)$ is the standard error of the estimated CCC.

2. Replacing the left-censored data by the LLD method refers to the use of the CCC after replacing all non-detectable data by the applicable detection limit. The 95% confidence interval (CI) of this method is calculated by using $\hat{\rho}_c \pm Z_{0.025}SE(\hat{\rho}_c)$ where $\hat{\rho}_c$ is the estimated CCC, $Z_{0.025}$ is the critical value of the standard normal distribution, and $SE(\hat{\rho}_c)$ is the standard error of the estimated CCC.

3. Replacing the left-censored data by one-half of the LLD method refers to the calculation of the CCC using all pairs after replacing the non-detectable data with 0.5 times the detection limit. The 95% confidence interval (CI) of this method is calculated by using $\hat{\rho}_c \pm Z_{0.025}SE(\hat{\rho}_c)$ where $\hat{\rho}_c$ is the estimated CCC, $Z_{0.025}$ is the critical value of the standard normal distribution, and $SE(\hat{\rho}_c)$ is the standard error of the estimated CCC.

4. Replacing the left-censored data by $c \times$ LLD method refers to the situation in which we first generate a random number from the uniform [0, 1] distribution, say $c$. Then, we replace each non-detectable data point with $c$ times the detection limit. A new value of $c$ is determined for each non-detectable data point. The 95% confidence interval (CI) of this method is calculated by using $\hat{\rho}_c \pm Z_{0.025}SE(\hat{\rho}_c)$ where $\hat{\rho}_c$ is the estimated CCC, $Z_{0.025}$ is the critical value of the standard normal distribution, and $SE(\hat{\rho}_c)$ is the standard error of the estimated CCC.
5. The ML approach is performed by constructing a likelihood function based on the bivariate normal distribution of the data in the detectable range, and then extrapolating into the region below the LLD. The 95% confidence interval (CI) of this method is calculated by using $\hat{\rho}_c \pm Z_{0.025}SE(\hat{\rho}_c)$ where $\hat{\rho}_c$ is the estimated CCC, $Z_{0.025}$ is the critical value of the standard normal distribution, and $SE(\hat{\rho}_c)$ is the standard error of the estimated CCC.

### 3.3 Results

#### 3.3.1 Simulation

In Tables 3.1, 3.2, 3.3, 3.4, 3.5 and 3.6, we report the results of a simulation study to assess the means and the standard deviations for estimating the CCC based on the ML approach and compare them to the four different methods that are described in the methods section, which frequently are applied in clinical research. In addition to means and standard deviations, we also report the relative bias, the mean of the standard error, and the percentage of 95% confidence intervals (CI) that include the true value of CCC for the 1,000 simulated data sets.
Table 3.1. Simulation Results Based on 1000 Data Sets with Sample Size of 100 – Percent Censoring (25%, 25%)

<table>
<thead>
<tr>
<th>True $\rho$</th>
<th>True $\rho_c$</th>
<th>Method</th>
<th>Mean $\hat{\rho}_c$</th>
<th>Relative bias (%)</th>
<th>Empirical Mean SD</th>
<th>Empirical Mean SE</th>
<th>The percentage of 95% confidence intervals that include true value of CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.238</td>
<td>1. Delete the pair</td>
<td>0.1405</td>
<td>-40.97</td>
<td>0.1119</td>
<td>0.1087</td>
<td>85.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Replace by LOD</td>
<td>0.2072</td>
<td>-12.94</td>
<td>0.1214</td>
<td>0.0856</td>
<td>91.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Replace by $0.5 \times$ LOD</td>
<td>0.1863</td>
<td>-21.72</td>
<td>0.1630</td>
<td>0.0848</td>
<td>88.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Replace by $c \times$ LOD</td>
<td>0.1951</td>
<td>-18.03</td>
<td>0.1240</td>
<td>0.0851</td>
<td>89.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. ML</td>
<td>0.2346</td>
<td>-1.43</td>
<td>0.0944</td>
<td>0.0939</td>
<td>94.3</td>
</tr>
<tr>
<td>0.50</td>
<td>0.476</td>
<td>1. Delete the pair</td>
<td>0.2966</td>
<td>-37.69</td>
<td>0.1197</td>
<td>0.0999</td>
<td>61.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Replace by LOD</td>
<td>0.4169</td>
<td>-12.42</td>
<td>0.1569</td>
<td>0.0735</td>
<td>86.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Replace by $0.5 \times$ LOD</td>
<td>0.3664</td>
<td>-23.03</td>
<td>0.3516</td>
<td>0.0742</td>
<td>78.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Replace by $c \times$ LOD</td>
<td>0.3821</td>
<td>-19.73</td>
<td>0.1541</td>
<td>0.0744</td>
<td>78.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. ML</td>
<td>0.4701</td>
<td>-1.24</td>
<td>0.0807</td>
<td>0.0780</td>
<td>93.6</td>
</tr>
<tr>
<td>0.75</td>
<td>0.714</td>
<td>1. Delete the pair</td>
<td>0.5307</td>
<td>-25.67</td>
<td>0.0886</td>
<td>0.0762</td>
<td>38.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Replace by LOD</td>
<td>0.6582</td>
<td>-7.82</td>
<td>0.0996</td>
<td>0.0497</td>
<td>85.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Replace by $0.5 \times$ LOD</td>
<td>0.6140</td>
<td>-14.01</td>
<td>0.1482</td>
<td>0.0519</td>
<td>71.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Replace by $c \times$ LOD</td>
<td>0.6140</td>
<td>-14.01</td>
<td>0.1662</td>
<td>0.0521</td>
<td>72.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. ML</td>
<td>0.7077</td>
<td>-0.88</td>
<td>0.0525</td>
<td>0.0503</td>
<td>94.8</td>
</tr>
</tbody>
</table>
Table 3.2. Simulation Results Based on 1000 Data Sets with Sample Size of 100 – Percent Censoring (40%, 25%)

<table>
<thead>
<tr>
<th>True $\rho$</th>
<th>True $\rho_c$</th>
<th>Method</th>
<th>Mean $\hat{\rho}_c$</th>
<th>Relative bias (%)</th>
<th>Empirical Mean $\bar{\rho}_c$</th>
<th>SE $\text{Emp}$</th>
<th>The percentage of 95% confidence intervals that include true value of CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.238</td>
<td>1. Delete the pair</td>
<td>0.1240</td>
<td>-47.90</td>
<td>0.1279</td>
<td>0.1230</td>
<td>84.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Replace by LOD</td>
<td>0.1929</td>
<td>-18.95</td>
<td>0.1319</td>
<td>0.0843</td>
<td>86.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Replace by $0.5 \times$ LOD</td>
<td>0.1856</td>
<td>-22.02</td>
<td>0.1092</td>
<td>0.0845</td>
<td>85.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Replace by $c \times$ LOD</td>
<td>0.1850</td>
<td>-22.27</td>
<td>0.1180</td>
<td>0.0844</td>
<td>85.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. ML</td>
<td>0.2340</td>
<td>-1.68</td>
<td>0.0984</td>
<td>0.0971</td>
<td>93.9</td>
</tr>
<tr>
<td>0.50</td>
<td>0.476</td>
<td>1. Delete the pair</td>
<td>0.2754</td>
<td>-42.14</td>
<td>0.1190</td>
<td>0.1108</td>
<td>60.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Replace by LOD</td>
<td>0.3901</td>
<td>-18.05</td>
<td>0.1219</td>
<td>0.0727</td>
<td>80.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Replace by $0.5 \times$ LOD</td>
<td>0.3661</td>
<td>-23.09</td>
<td>0.1488</td>
<td>0.0736</td>
<td>73.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Replace by $c \times$ LOD</td>
<td>0.3664</td>
<td>-23.03</td>
<td>0.1603</td>
<td>0.0735</td>
<td>74.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. ML</td>
<td>0.4692</td>
<td>-1.43</td>
<td>0.0847</td>
<td>0.0811</td>
<td>93.8</td>
</tr>
<tr>
<td>0.75</td>
<td>0.714</td>
<td>1. Delete the pair</td>
<td>0.4966</td>
<td>-30.45</td>
<td>0.0978</td>
<td>0.0851</td>
<td>30.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Replace by LOD</td>
<td>0.6268</td>
<td>-12.21</td>
<td>0.1054</td>
<td>0.0500</td>
<td>68.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Replace by $0.5 \times$ LOD</td>
<td>0.6055</td>
<td>-15.20</td>
<td>0.1389</td>
<td>0.0512</td>
<td>64.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Replace by $c \times$ LOD</td>
<td>0.6083</td>
<td>-14.80</td>
<td>0.1306</td>
<td>0.0513</td>
<td>65.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. ML</td>
<td>0.7066</td>
<td>-1.04</td>
<td>0.0554</td>
<td>0.0528</td>
<td>94.8</td>
</tr>
</tbody>
</table>
### Table 3.3. Simulation Results Based on 1000 Data Sets with Sample Size of 50 – Percent Censoring (25%, 25%)

<table>
<thead>
<tr>
<th>True $\rho$</th>
<th>True $\rho_c$</th>
<th>Method</th>
<th>Mean $\hat{\rho}_c$</th>
<th>Relative bias (%)</th>
<th>Empirical Mean SD</th>
<th>Empirical Mean SE</th>
<th>The percentage of 95% confidence intervals that include true value of CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.238</td>
<td>1. Delete the pair</td>
<td>0.1402</td>
<td>-41.09</td>
<td>0.1693</td>
<td>0.1486</td>
<td>87.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Replace by LOD</td>
<td>0.2079</td>
<td>-12.65</td>
<td>0.1509</td>
<td>0.1172</td>
<td>88.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Replace by $0.5 \times$ LOD</td>
<td>0.1811</td>
<td>-23.91</td>
<td>0.2088</td>
<td>0.1163</td>
<td>83.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Replace by $c \times$ LOD</td>
<td>0.1949</td>
<td>-18.09</td>
<td>0.1813</td>
<td>0.1162</td>
<td>83.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. ML</td>
<td>0.2310</td>
<td>-2.94</td>
<td>0.1351</td>
<td>0.1304</td>
<td>92.7</td>
</tr>
<tr>
<td>0.50</td>
<td>0.476</td>
<td>1. Delete the pair</td>
<td>0.2936</td>
<td>-38.32</td>
<td>0.1481</td>
<td>0.1386</td>
<td>79.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Replace by LOD</td>
<td>0.4144</td>
<td>-12.94</td>
<td>0.4115</td>
<td>0.1016</td>
<td>87.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Replace by $0.5 \times$ LOD</td>
<td>0.3661</td>
<td>-23.09</td>
<td>0.1854</td>
<td>0.1024</td>
<td>79.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Replace by $c \times$ LOD</td>
<td>0.3722</td>
<td>-21.81</td>
<td>0.1770</td>
<td>0.1026</td>
<td>81.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. ML</td>
<td>0.4636</td>
<td>-2.61</td>
<td>0.1151</td>
<td>0.1095</td>
<td>93.4</td>
</tr>
<tr>
<td>0.75</td>
<td>0.714</td>
<td>1. Delete the pair</td>
<td>0.5169</td>
<td>-27.61</td>
<td>0.1341</td>
<td>0.1071</td>
<td>66.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Replace by LOD</td>
<td>0.6429</td>
<td>-9.96</td>
<td>0.1438</td>
<td>0.0694</td>
<td>88.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Replace by $0.5 \times$ LOD</td>
<td>0.5812</td>
<td>-18.60</td>
<td>0.2127</td>
<td>0.0720</td>
<td>78.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Replace by $c \times$ LOD</td>
<td>0.5909</td>
<td>-17.24</td>
<td>0.1917</td>
<td>0.0727</td>
<td>79.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. ML</td>
<td>0.7025</td>
<td>-1.61</td>
<td>0.0759</td>
<td>0.0716</td>
<td>93.6</td>
</tr>
<tr>
<td>True (\rho)</td>
<td>True (\rho_c)</td>
<td>Method</td>
<td>Mean (\hat{\rho}_c)</td>
<td>Relative bias (%)</td>
<td>Empirical Mean SD</td>
<td>Empirical Mean SE</td>
<td>The percentage of 95% confidence intervals that include true value of CCC</td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>-------------------------</td>
<td>------------------------</td>
<td>---------------------</td>
<td>-------------------</td>
<td>------------------</td>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>0.25</td>
<td>0.238</td>
<td>1. Delete the pair</td>
<td>0.1334</td>
<td>-43.95</td>
<td>0.1858</td>
<td>0.1652</td>
<td>88.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Replace by LOD</td>
<td>0.1908</td>
<td>-19.83</td>
<td>0.1525</td>
<td>0.1150</td>
<td>84.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Replace by (0.5 \times) LOD</td>
<td>0.1796</td>
<td>-24.54</td>
<td>0.1467</td>
<td>0.1155</td>
<td>80.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Replace by (c \times) LOD</td>
<td>0.1780</td>
<td>-25.22</td>
<td>0.1528</td>
<td>0.1152</td>
<td>81.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. ML</td>
<td>0.2296</td>
<td>-3.53</td>
<td>0.1395</td>
<td>0.1345</td>
<td>92.4</td>
</tr>
<tr>
<td>0.50</td>
<td>0.476</td>
<td>1. Delete the pair</td>
<td>0.2735</td>
<td>-42.54</td>
<td>0.1693</td>
<td>0.1516</td>
<td>76.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Replace by LOD</td>
<td>0.3673</td>
<td>-22.84</td>
<td>0.2110</td>
<td>0.1004</td>
<td>80.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Replace by (0.5 \times) LOD</td>
<td>0.3366</td>
<td>-29.29</td>
<td>0.2771</td>
<td>0.1011</td>
<td>74.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Replace by (c \times) LOD</td>
<td>0.3537</td>
<td>-25.70</td>
<td>0.1788</td>
<td>0.1011</td>
<td>75.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. ML</td>
<td>0.4617</td>
<td>-3.00</td>
<td>0.1206</td>
<td>0.1135</td>
<td>93.2</td>
</tr>
<tr>
<td>0.75</td>
<td>0.714</td>
<td>1. Delete the pair</td>
<td>0.4913</td>
<td>-31.19</td>
<td>0.1373</td>
<td>0.1184</td>
<td>62.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Replace by LOD</td>
<td>0.6022</td>
<td>-15.66</td>
<td>0.1705</td>
<td>0.0696</td>
<td>78.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Replace by (0.5 \times) LOD</td>
<td>0.5706</td>
<td>-20.08</td>
<td>0.2039</td>
<td>0.0710</td>
<td>74.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Replace by (c \times) LOD</td>
<td>0.5689</td>
<td>-20.32</td>
<td>0.2151</td>
<td>0.0713</td>
<td>74.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. ML</td>
<td>0.7003</td>
<td>-1.92</td>
<td>0.0805</td>
<td>0.0750</td>
<td>93.2</td>
</tr>
</tbody>
</table>
### Table 3.5. Simulation Results Based on 1000 Data Sets with Sample Size of 25 – Percent Censoring (25%, 25%)

<table>
<thead>
<tr>
<th>True $\rho$</th>
<th>True $\rho_c$</th>
<th>Method</th>
<th>Mean $\hat{\rho}_c$</th>
<th>Relative bias (%)</th>
<th>Empirical SD</th>
<th>Empirical SE</th>
<th>Percentage of 95% confidence intervals that include true value of CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.238</td>
<td>1. Delete the pair</td>
<td>0.1300</td>
<td>-45.38</td>
<td>0.2287</td>
<td>0.2002</td>
<td>85.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Replace by LOD</td>
<td>0.2159</td>
<td>-9.29</td>
<td>0.2176</td>
<td>0.1563</td>
<td>84.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Replace by 0.5 × LOD</td>
<td>0.1821</td>
<td>-23.49</td>
<td>0.2280</td>
<td>0.1550</td>
<td>78.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Replace by $c \times LOD$</td>
<td>0.1897</td>
<td>-20.29</td>
<td>0.2215</td>
<td>0.1552</td>
<td>78.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. ML</td>
<td>0.2225</td>
<td>-6.51</td>
<td>0.1905</td>
<td>0.1783</td>
<td>93.1</td>
</tr>
<tr>
<td>0.50</td>
<td>0.476</td>
<td>1. Delete the pair</td>
<td>0.2862</td>
<td>-39.87</td>
<td>0.2088</td>
<td>0.1853</td>
<td>83.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Replace by LOD</td>
<td>0.3963</td>
<td>-16.74</td>
<td>0.2030</td>
<td>0.1356</td>
<td>83.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Replace by 0.5 × LOD</td>
<td>0.3538</td>
<td>-25.67</td>
<td>0.2277</td>
<td>0.1372</td>
<td>76.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Replace by $c \times LOD$</td>
<td>0.3642</td>
<td>-23.49</td>
<td>0.2105</td>
<td>0.1372</td>
<td>78.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. ML</td>
<td>0.4496</td>
<td>-5.55</td>
<td>0.1650</td>
<td>0.1526</td>
<td>93.3</td>
</tr>
<tr>
<td>0.75</td>
<td>0.714</td>
<td>1. Delete the pair</td>
<td>0.4850</td>
<td>-32.07</td>
<td>0.1951</td>
<td>0.1489</td>
<td>77.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Replace by LOD</td>
<td>0.6031</td>
<td>-15.53</td>
<td>0.2122</td>
<td>0.0960</td>
<td>83.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Replace by 0.5 × LOD</td>
<td>0.5467</td>
<td>-23.43</td>
<td>0.2545</td>
<td>0.0984</td>
<td>77.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Replace by $c \times LOD$</td>
<td>0.5436</td>
<td>-23.87</td>
<td>0.2700</td>
<td>0.0991</td>
<td>77.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. ML</td>
<td>0.6897</td>
<td>-3.40</td>
<td>0.1136</td>
<td>0.1027</td>
<td>92.1</td>
</tr>
</tbody>
</table>
Table 3.6. Simulation Results Based on 1000 Data Sets with Sample Size of 25 – Percent Censoring (40%, 25%)

<table>
<thead>
<tr>
<th>True ρ</th>
<th>True ρc</th>
<th>Method</th>
<th>Mean  $\hat{\rho}_c$</th>
<th>Relative bias (%)</th>
<th>Empirical Mean SD</th>
<th>Empirical Mean SE</th>
<th>The percentage of 95% confidence intervals that include true value of CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.238</td>
<td>1. Delete the pair</td>
<td>0.1357</td>
<td>-42.98</td>
<td>0.2485</td>
<td>0.2130</td>
<td>85.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Replace by LOD</td>
<td>0.1703</td>
<td>-28.45</td>
<td>0.4223</td>
<td>0.1530</td>
<td>80.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Replace by 0.5 × LOD</td>
<td>0.1822</td>
<td>-23.45</td>
<td>0.2936</td>
<td>0.1535</td>
<td>74.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Replace by $c \times$ LOD</td>
<td>0.1751</td>
<td>-26.45</td>
<td>0.2033</td>
<td>0.1533</td>
<td>75.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. ML</td>
<td>0.2193</td>
<td>-7.86</td>
<td>0.1948</td>
<td>0.1824</td>
<td>93.0</td>
</tr>
<tr>
<td>0.50</td>
<td>0.476</td>
<td>1. Delete the pair</td>
<td>0.2667</td>
<td>-43.97</td>
<td>0.2296</td>
<td>0.1982</td>
<td>82.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Replace by LOD</td>
<td>0.3628</td>
<td>-23.78</td>
<td>0.2088</td>
<td>0.1334</td>
<td>78.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Replace by 0.5 × LOD</td>
<td>0.3307</td>
<td>-30.53</td>
<td>0.2283</td>
<td>0.1350</td>
<td>71.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Replace by $c \times$ LOD</td>
<td>0.3304</td>
<td>-30.59</td>
<td>0.2416</td>
<td>0.1352</td>
<td>72.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. ML</td>
<td>0.4449</td>
<td>-6.53</td>
<td>0.1693</td>
<td>0.1570</td>
<td>93.3</td>
</tr>
<tr>
<td>0.75</td>
<td>0.714</td>
<td>1. Delete the pair</td>
<td>0.4600</td>
<td>-35.57</td>
<td>0.2050</td>
<td>0.1608</td>
<td>75.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Replace by LOD</td>
<td>0.5599</td>
<td>-21.58</td>
<td>0.2482</td>
<td>0.0947</td>
<td>78.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Replace by 0.5 × LOD</td>
<td>0.5246</td>
<td>-26.53</td>
<td>0.2619</td>
<td>0.0961</td>
<td>72.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Replace by $c \times$ LOD</td>
<td>0.5266</td>
<td>-26.25</td>
<td>0.2678</td>
<td>0.0964</td>
<td>74.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. ML</td>
<td>0.6847</td>
<td>-4.10</td>
<td>0.1180</td>
<td>0.1074</td>
<td>92.2</td>
</tr>
</tbody>
</table>
As demonstrated in Tables 3.1, and 3.2, the estimates from the four simple approaches are obviously biased, although the replacement of non-detectable data by a fraction of the detection limit or the detection limit itself is clearly preferable to discarding the pair method for all range of sample sizes. From Table 3.1 (sample size of 100 paired data points, 25% left-censoring for X, 25% left-censoring for Y, and a true CCC of 0.238), the relative bias is -1.43% for the ML method, -40.97% for the data deletion method, and it ranges between -12.94% and -21.72% for the simple data imputation methods. These four ad hoc methods also overstate the precision by underestimating the standard error set because the same value is imputed repeatedly. As expected, the ML method provides an excellent estimate of the true value of the CCC even when the censoring percentages increased, but it tends to slightly underestimate the true value. Moreover, the ML approach yields the smallest relative bias and the highest percentage of the 95% CI that include the true value of CCC among five methods. To see the impact of the percent of censoring, in Table 3.2, we increase the censoring rate to 40%. The relative biases are increased in all approaches. However, the ML approach still yields the smallest relative bias. From Table 3.2 (sample size of 100 paired data points, 40% left-censoring for X, 25% left-censoring for Y, and a true CCC of 0.238), the relative bias is -1.68% for the ML method, -47.90% for the data deletion method, and it ranges between -18.95% and -22.27% for the simple data imputation methods. The ML method also has the highest percentage of confidence intervals that include the true value of the CCC.

Due to the large sample size of both assays (sample size =100) in Tables 3.1 and 3.2, the ML method displays an excellent result for estimating the CCC with respect to the relative bias and the percentage of confidence intervals that include the true value of the CCC. However, if the sample size were smaller, then the ML method might produce less convincing results. To illustrate this point, we re-conduct the simulation studies with sample sizes = 50 (Tables 3.3 and 3.4) and sample sizes = 25 (Tables 3.5 and 3.6). In both cases, the ML method still performs best among the five approaches according to the means and the standard deviations for estimating the CCC, the relative bias, mean of the standard error, and the percentage of 95% CI that include the true value of CCC.
3.3.2 Example

We illustrate these issues further via a urine stability study to assess agreement for two assays with lower limits of detection. The data set came from the multi-center ASSESS AKI Study (the Assessment, Serial Evaluation, and Subsequent Sequelae of Acute Kidney Injury). The data set was originally analyzed by Parikh et al. [2]. The purpose of the ASSESS AKI sub-study was to determine the agreement between the measurements of the urinary biomarkers collected under a standard condition and under different experimental conditions, denoted as Process A, Process B, and Process C. Each experimental situation consisted of 50 paired samples (a selected process versus the standard). There are two biomarkers that we consider here: urine Interleukin 18 (IL-18; LLD = 12.5 pg/ml), and urine Cystatin C (LLD = 0.005 mg/ml). The IL-18 contained 99 undetectable readings (out of a total of 300), yielding a 33% left-censoring rate. The Cystatin C contained 80 undetectable readings, for a 26.7% left-censoring rate. A natural logarithm transformation was applied to both the IL-18 and Cystatin C readings. We treat the natural logarithm of Process A, Process B, and Process C as the X variable and the natural logarithm of the reference standard as the Y variable.

Tables 3.7 and 3.8 summarize the results based on the four ad hoc approaches and the ML method, for estimating the CCC when comparing the reference standard process to Process A, Process B, and Process C for IL-18 and Cystatin C, respectively. As this example suggests, the four simple approaches can lead to CCC estimates that are different than the CCC estimated from the ML method. For example, from comparing Process B to the standard for urine IL-18 in Table 3.7, the CCC estimate is 0.73 from the data deletion method, 0.61 from each of the simple data imputation methods, and 0.68 from the ML method.
Table 3.7. Concordance Correlation Coefficients (and 95% Confidence Intervals) for 3 Processes Using 5 Methods Based on IL-18 Assay

<table>
<thead>
<tr>
<th>Processes</th>
<th>Method 1</th>
<th>Method 2</th>
<th>Method 3</th>
<th>Method 4</th>
<th>Method 5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Delete pair</td>
<td>Replace by LOD</td>
<td>Replace by 0.5×LOD</td>
<td>Replace by c×LOD</td>
<td>Maximum Likelihood</td>
</tr>
<tr>
<td>A (Initial 48 hours: 4°C vs −80°C)</td>
<td>0.8801</td>
<td>0.8228</td>
<td>0.8228</td>
<td>0.8228</td>
<td>0.8314</td>
</tr>
<tr>
<td></td>
<td>(0.81, 0.95)</td>
<td>(0.73, 0.91)</td>
<td>(0.73, 0.91)</td>
<td>(0.73, 0.91)</td>
<td>(0.74, 0.92)</td>
</tr>
<tr>
<td>B (Initial 48 hours: 25°C vs −80°C)</td>
<td>0.7344</td>
<td>0.6081</td>
<td>0.6081</td>
<td>0.6081</td>
<td>0.6819</td>
</tr>
<tr>
<td></td>
<td>(0.56, 0.91)</td>
<td>(0.43, 0.77)</td>
<td>(0.43, 0.77)</td>
<td>(0.43, 0.77)</td>
<td>(0.51, 0.85)</td>
</tr>
<tr>
<td>C (Centrifuge vs No Centrifuge)</td>
<td>0.9886</td>
<td>0.9896</td>
<td>0.9896</td>
<td>0.9896</td>
<td>0.9876</td>
</tr>
<tr>
<td></td>
<td>(0.98, 1.00)</td>
<td>(0.98, 1.00)</td>
<td>(0.98, 1.00)</td>
<td>(0.98, 1.00)</td>
<td>(0.99, 1.00)</td>
</tr>
</tbody>
</table>
Table 3.8. Concordance Correlation Coefficients (and 95% Confidence Intervals) for 3 Processes Using 5 Methods Based on Cystatin-C Assay

<table>
<thead>
<tr>
<th>Processes</th>
<th>Method 1 Delete pair</th>
<th>Method 2 Replace by LOD</th>
<th>Method 3 Replace by 0.5×LOD</th>
<th>Method 4 Replace by c×LOD</th>
<th>Method 5 Maximum Likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>A (Initial 48 hours:</td>
<td>0.9348</td>
<td>0.9641</td>
<td>0.9641</td>
<td>0.9641</td>
<td>0.9735</td>
</tr>
<tr>
<td>4°C vs −80°C</td>
<td>(0.89, 0.98)</td>
<td>(0.94, 0.98)</td>
<td>(0.94, 0.98)</td>
<td>(0.94, 0.98)</td>
<td>(0.95, 0.99)</td>
</tr>
<tr>
<td>B (Initial 48 hours:</td>
<td>0.9320</td>
<td>0.9514</td>
<td>0.9514</td>
<td>0.9514</td>
<td>0.9471</td>
</tr>
<tr>
<td>25°C vs −80°C</td>
<td>(0.89, 0.98)</td>
<td>(0.92, 0.98)</td>
<td>(0.92, 0.98)</td>
<td>(0.92, 0.98)</td>
<td>(0.91, 0.98)</td>
</tr>
<tr>
<td>C (Centrifuge vs No</td>
<td>0.9985</td>
<td>0.9982</td>
<td>0.9758</td>
<td>0.9653</td>
<td>0.9999</td>
</tr>
<tr>
<td>Centrifuge)</td>
<td>(0.99, 1.00)</td>
<td>(0.99, 1.00)</td>
<td>(0.96, 0.99)</td>
<td>(0.95, 0.98)</td>
<td>(0.99, 1.00)</td>
</tr>
</tbody>
</table>
3.4 Discussion

Biomarkers are being discovered at an accelerated rate due to availability of genomic and proteomic technologies [1]. Several of these candidate biomarkers are undergoing validation to diagnose diseases and serve as indices for predicting health outcomes. The main purpose of our study was to assist the biomarker development program by confirming that the simple data imputation approaches and the deletion of data are not optimal techniques for arriving at accurate (unbiased) results with the appropriate level of precision in the presence of left-censored data.

Many researchers have stressed the importance of data that are below the LLD [2,4–6]. Hornung and Reed [6] proposed three methods of estimation with a left-censored lognormal distribution: a maximum likelihood (ML) method and two methods involving the limit of detection. However, they conclude that the ML method is complex to calculate, so they recommend using the one-half of the LLD. Lyles et al. [4] evaluated the Pearson’s correlation coefficient when a subset of data points was below the LLD by using the ML approach under the assumption of bivariate normality. They showed that the ML method was the most accurate among the proposed methods. Barnhart et al. [5] presented a generalized estimating equations (GEE) approach for estimating parameters to calculate the concordance correlation coefficient (CCC) [3], which is a measure of agreement ranging between -1 and +1 for paired data. The GEE approach works well and does not require the bivariate normality assumption if the sample size is large enough, and it is comparable to the ML approach when the bivariate normality assumption is appropriate. Parikh et al. [2] performed a prospective study on hospitalized patients with almost 60% of patients having acute kidney injury (AKI). Five urine biomarkers were used to compare the stability of short-term storage and processing by using the CCC as a measure of agreement. To estimate the CCC, the authors applied the ML method using log-transformed data and accounting for values below the LLD.

We have illustrated with our computer simulation study that the estimation of the CCC from the imputation methods or data deletion lead to biased estimates compared to the ML approach. We also have shown via the computer simulation study that the proportion of left-censored data significantly impacts the degree of bias in estimating the CCC. Our simulation study shows that the ML approach based on the bivariate normality assumption works best among all of the studied approaches.
The advantages of the ML approach are that it is accurate (small relative bias) and accounts for the variability in the data set appropriately. Additionally, it uses all the available data for the statistical analysis, in contrast to the data deletion approach that only uses sample pairs with both values above the LLD in the analysis. The estimates from the data deletion approach are obviously biased and result in a (1) large relative bias and (2) a high value of the standard error due to a small sample size from deleting paired data points. Although assigning a fixed value such as the LLD (or one-half of the LLD or the multiplication of the LLD by a random number from the uniform(0,1) distribution), yields smaller relative biases compared to the data deletion approach, the precision from these methods is overestimated due to the assignment of the same value to data below the LLD.

Although we did not investigate the performance of the ML method for censoring above 40%, we expect that the ML method still will perform well when censoring exceeds 50%. Lyles et al. [4] investigated 60% censoring for their situation and the ML method still maintained a high level of accuracy.

3.5 Conclusions

The ML approach is very accurate in that it yields small relative biases if the assumption of bivariate normality is appropriate, and it can be readily implemented using SAS PROC NLMIXED. Thus, our simulation study suggests that the ML approach is best for biomarker assay development where paired results need to be compared.
Chapter 4  
New Class of Weibull Distributions

4.1 Bivariate Survival Functions

Let

\[ X = \begin{bmatrix} X_1 & X_2 \end{bmatrix}^T \]

denote a random vector such that each \( X_i \) is a nonnegative-valued random variable, \( i = 1, 2 \). Let

- \( f_i(x_i) \) denote the density function
- \( F_i(x_i) = \int_0^{x_i} f_i(u) \, du \) denote the cumulative distribution function
- \( S_i(x_i) = 1 - F_i(x_i) = \int_{x_i}^{\infty} f_i(u) \, du \) denote the survival function
- \( h_i(x_i) = \frac{f_i(x_i)}{S_i(x_i)} \) denote the hazard function
- \( H_i(x_i) = \int_0^{x_i} h_i(u) \, du \) denote the cumulative hazard function

for the random variable \( X_i, i = 1, 2 \). All of these functions are nonnegative-valued, \( S_i(x_i) \) is a decreasing function of \( x_i \), and \( H_i(x_i) \) is an increasing function of \( x_i, i = 1, 2 \).

For \( i = 1, 2 \), let \( g_i(u) \) denote a real-valued function of \( u, u \geq 0 \), such that

- \( g_i(0) = 0 \)
- \( 0 \leq g_i(u) \leq 1 \)
• $g_i(u)$ is increasing in $u$, i.e., $\frac{\partial}{\partial u} g_i(u) = g'_i(u) > 0$ for $u > 0$

• $\frac{\partial}{\partial u} g_i(u) = g'_i(u) \leq 1$

The following are examples of such functions:

• $g(u) = \frac{u}{(1+u)}$ with $g'(u) = \frac{1}{(1+u)^2}, u \geq 0$

• $g(u) = \frac{2}{\pi} \tan^{-1}(u)$ with $g'(u) = \frac{2}{\pi(1+u^2)}, u \geq 0$

• $g(u) = \frac{1-e^{-u}}{1+e^{-u}}$ with $g'(u) = \frac{e^{-2u}}{(1+e^{-u})^2}, u \geq 0$

• $g(u) = \frac{\log(1+u)}{1+\log(1+u)}$ with $g'(u) = \frac{1}{(1+u)(1+\log(1+u))}, u \geq 0$

• $g(u) = 1 - e^{-u}$ with $g'(u) = e^{-u}, u \geq 0$

• $g(u) = 1 - (1+u)e^{-u}$ with $g'(u) = ue^{-u}, u \geq 0$

• $g(u) = 1 - e^{-\frac{u^2}{2}}$ with $g'(u) = ue^{-\frac{u^2}{2}}, u \geq 0$

Construct a bivariate survival function as

$$S_{12}(x_1, x_2) = S_1(x_1)S_2(x_2) \left[ 1 + \alpha_{12}g_1(H_1(x_1))g_2(H_2(x_2)) \right] \quad (4.1)$$

where $\alpha_{12}$ is an association or correlation parameter, $-1 \leq \alpha_{12} \leq +1$. Notice that $\alpha_{12} = 0$ yields the independence of $X_1$ and $X_2$. Also, notice that $S_{12}(x_1, 0) = S_1(x_1)$ and $S_{12}(0, x_2) = S_2(x_2)$, so that the marginal survival functions coincide with the original univariate survival functions.

Model (4.1) for the bivariate survival function is similar in construction to the Farlie-Gumbel-Morgenstern family of bivariate cumulative distribution functions defined as

$$F_{12}(x_1, x_2) = F_1(x_1)F_2(x_2) \left[ 1 + \alpha_{12}S_1(x_1)S_2(x_2) \right]$$

which can be re-expressed as a bivariate survival function as

$$S_{12}(x_1, x_2) = S_1(x_1)S_2(x_2) \left[ 1 + \alpha_{12}F_1(x_1)F_2(x_2) \right]$$

The distinction between the Farlie-Gumbel-Morgenstern bivariate survival function and the bivariate survival function in Model (4.1) is that the former is based solely on
the marginal survival functions, whereas the latter is based on the marginal survival functions and the marginal cumulative hazard functions.

Next, we determine the first partial derivatives of $S_{12}(x_1, x_2)$ for Model (4.1):

$$\frac{\partial}{\partial x_1} S_{12}(x_1, x_2) = \frac{\partial}{\partial x_1} S_1(x_1)S_2(x_2) \left[ 1 + \alpha_{12} g_1(H_1(x_1)) g_2(H_2(x_2)) \right]$$

$$= S_1(x_1)S_2(x_2) \left[ \alpha_{12} g'_1(H_1(x_1)) h_1(x_1) g_2(H_2(x_2)) \right]$$

$$+ \left[ 1 + \alpha_{12} g_1(H_1(x_1)) g_2(H_2(x_2)) S_2(x_2)(-f_1(x_1)) \right]$$

$$= \frac{f_1(x_1)}{h_1(x_1)} S_2(x_2) \left[ \alpha_{12} g'_1(H_1(x_1)) h_1(x_1) g_2(H_2(x_2)) \right]$$

$$+ \left[ 1 + \alpha_{12} g_1(H_1(x_1)) g_2(H_2(x_2)) S_2(x_2)(-f_1(x_1)) \right]$$

$$= -f_1(x_1) S_2(x_2) \left[ -\alpha_{12} g'_1(H_1(x_1)) g_2(H_2(x_2)) + 1 + \alpha_{12} g_1(H_1(x_1)) g_2(H_2(x_2)) \right]$$

$$= -f_1(x_1) S_2(x_2) \left[ 1 + \alpha_{12} g_2(H_2(x_2)) \left( g_1(H_1(x_1)) - g'_1(H_1(x_1)) \right) \right]$$

Likewise,

$$\frac{\partial}{\partial x_2} S_{12}(x_1, x_2) = -S_1(x_1) f_2(x_2) \left[ 1 + \alpha_{12} g_1(H_1(x_1)) \left( g_2(H_2(x_2)) - g'_2(H_2(x_2)) \right) \right]$$

The first partial derivatives are negative, indicating that $S_{12}(x_1, x_2)$ is non-increasing in $x_1$ and non-increasing in $x_2$. The second partial derivative of $S_{12}(x_1, x_2)$ for Model (4.1) with respect to $x_1$ and $x_2$ yields the bivariate density function as
\[ f_{12}(x_1, x_2) = (-1)^2 \frac{\partial^2}{\partial x_2 \partial x_1} S_{12}(x_1, x_2) \]

\[ = \frac{\partial}{\partial x_2} \left\{ -f_1(x_1)S_2(x_2) \left[ 1 + \alpha_{12}g_2(H_2(x_2)) \left( g_1(H_1(x_1)) - g'_1(H_1(x_1)) \right) \right] \right\} \]

\[ = \left( -f_1(x_1)S_2(x_2) \right) \left[ \left( g_1(H_1(x_1)) - g'_1(H_1(x_1)) \right) \alpha_{12}g'_2(H_2(x_2))h_2(x_2) \right] \]

\[ + \left[ 1 + \alpha_{12}g_2(H_2(x_2)) \left( g_1(H_1(x_1)) - g'_1(H_1(x_1)) \right) \right] (-f_1(x_1))(-f_2(x_2)) \]

\[ = \left( f_1(x_1)f_2(x_2) \right) \left[ -\alpha_{12}g'_2(H_2(x_2)) \left( g_1(H_1(x_1)) - g'_1(H_1(x_1)) \right) \right] + 1 \]

\[ + \alpha_{12}g_2(H_2(x_2)) \left( g_1(H_1(x_1)) - g'_1(H_1(x_1)) \right) \]

\[ = f_1(x_1)f_2(x_2) \left\{ 1 + \alpha_{12} \left( g_1(H_1(x_1)) - g'_1(H_1(x_1)) \right) \left( g_2(H_2(x_2)) - g'_2(H_2(x_2)) \right) \right\} \]

We have confirmed that \( S_{12}(x_1, x_2) \) is a bivariate survival function because we have demonstrated the following properties:

1. \( S_{12}(x_1, x_2) = 1 \)

2. \( \lim_{x_1 \to \infty} S_{12}(x_1, x_2) = \lim_{x_2 \to \infty} S_{12}(x_1, x_2) = 0 \)

3. \( \frac{\partial}{\partial x_1} S_{12}(x_1, x_2) \leq 0 \) and \( \frac{\partial}{\partial x_2} S_{12}(x_1, x_2) \leq 0 \)

4. \( \frac{\partial^2}{\partial x_2 \partial x_1} S_{12}(x_1, x_2) = f_{12}(x_1, x_2) \geq 0 \)

Another structure to consider for a bivariate survival function is

\[ S_{12}(x_1, x_2) = S_1(x_1)S_2(x_2) \left( 1 + g_1(H_1(x_1))g_2(H_2(x_2)) \right)^{\alpha_{12}} \]  (4.2)
Next, we determine the first partial derivatives of $S_{12}(x_1, x_2)$ for Model (3.2):

\[
\frac{\partial}{\partial x_1} S_{12}(x_1, x_2) = \frac{\partial}{\partial x_1} \left[ S_1(x_1)S_2(x_2) \left( 1 + g_1(H_1(x_1))g_2(H_2(x_2)) \right)^{\alpha_{12}} \right]
\]

\[
= S_1(x_1)S_2(x_2) \alpha_{12} \left( 1 + g_1(H_1(x_1))g_2(H_2(x_2)) \right)^{\alpha_{12}-1} g_1'(H_1(x_1))h_1(x_1)
\]

\[
+ \left[ \left( 1 + g_1(H_1(x_1))g_2(H_2(x_2)) \right)^{\alpha_{12}} S_2(x_2)(-f_1(x_1)) \right]
\]

\[
= \frac{f_1(x_1)}{h_1(x_1)} S_2(x_2) \alpha_{12} \left( 1 + g_1(H_1(x_1))g_2(H_2(x_2)) \right)^{\alpha_{12}-1} g_1'(H_1(x_1))h_1(x_1)
\]

\[
+ \left[ \left( 1 + g_1(H_1(x_1))g_2(H_2(x_2)) \right)^{\alpha_{12}} S_2(x_2)(-f_1(x_1)) \right]
\]

\[
= -f_1(x_1)S_2(x_2) \left( 1 + g_1(H_1(x_1))g_2(H_2(x_2)) \right)^{\alpha_{12}-1}
\]

\[
\times \left( 1 + \alpha_{12}g_1'(H_1(x_1)) + g_1(H_1(x_1))g_2(H_2(x_2)) \right)
\]

Likewise,

\[
\frac{\partial}{\partial x_2} S_{12}(x_1, x_2) = \frac{\partial}{\partial x_2} \left[ S_1(x_1)S_2(x_2) \left( 1 + g_1(H_1(x_1))g_2(H_2(x_2)) \right)^{\alpha_{12}} \right]
\]

\[
= -S_1(x_1)f_2(x_2) \left( 1 + g_1(H_1(x_1))g_2(H_2(x_2)) \right)^{\alpha_{12}-1}
\]

\[
\times \left( 1 + \alpha_{12}g_2'(H_2(x_2)) + g_1(H_1(x_1))g_2(H_2(x_2)) \right)
\]

The first partial derivatives are negative, indicating that $S_{12}(x_1, x_2)$ is non-increasing in $x_1$ and non-increasing in $x_2$. The second partial derivative of $S_{12}(x_1, x_2)$ for Model (3.2) with respect to $x_1$ and $x_2$ yields the bivariate density function as
\[ f_{12}(x_1, x_2) = (-1)^2 \frac{\partial^2}{\partial x_2 \partial x_1} S_{12}(x_1, x_2) \]

\[ = -f_1(x_1)S_2(x_2) \left( 1 + g_1(H_1(x_1))g_2(H_2(x_2)) \right)^{\alpha_{12}-1} g_1(H_1(x_1))g_2'(H_2(x_2))h_2(x_2) \]

\[ + \left( 1 + \alpha_{12}g_1'(H_1(x_1)) + g_1(H_1(x_1))g_2(H_2(x_2)) \right) \]

\[ \times \left[ -f_1(x_1)S_2(x_2)(\alpha_{12} - 1) \left( 1 + g_1(H_1(x_1))g_2(H_2(x_2)) \right)^{\rho_{12}-2} g_2'(H_2(x_2))h_2(x_2) \right. \]

\[ + \left. \left( 1 + g_1(H_1(x_1))g_2(H_2(x_2)) \right)^{\alpha_{12}-1} \right] \]

\[ = -f_1(x_1)f_2(x_2) \left( 1 + g_1(H_1(x_1))g_2(H_2(x_2)) \right)^{\alpha_{12}-1} g_1(H_1(x_1))g_2'(H_2(x_2)) \]

\[ + \left( 1 + \alpha_{12}g_1'(H_1(x_1)) + g_1(H_1(x_1))g_2(H_2(x_2))f_1(x_1)f_2(x_2) \right) \]

\[ \times \left( 1 + g_1(H_1(x_1))g_2(H_2(x_2)) \right)^{\alpha_{12}-2} \]

\[ \times \left[ - (\alpha_{12} - 1)g_2'(H_2(x_2)) + \left( 1 + g_1(H_1(x_1))g_2(H_2(x_2)) \right) \right] \]

\[ = f_1(x_1)f_2(x_2) \left( 1 + g_1(H_1(x_1))g_2(H_2(x_2)) \right)^{\alpha_{12}-2} \]

\[ \times \left[ \left( 1 + g_1(H_1(x_1))g_2(H_2(x_2)) \right)g_1(H_1(x_1))g_2'(H_2(x_2)) \right. \]

\[ + \left( 1 + \alpha_{12}g_1'(H_1(x_1)) + g_1(H_1(x_1))g_2(H_2(x_2)) \right) \]

\[ \times \left( 1 + g_1(H_1(x_1))g_2(H_2(x_2)) - (\alpha_{12} - 1)g_2'(H_2(x_2)) \right) \]

Similar to model (4.1), we have shown all four properties below to confirm that \( S_{12}(x_1, x_2) \) in model (4.2) is a bivariate survival function.

1. \( S_{12}(x_1, x_2) = 1 \)

2. \( \lim_{x_1 \to \infty} S_{12}(x_1, x_2) = \lim_{x_2 \to \infty} S_{12}(x_1, x_2) = 0 \)

3. \( \frac{\partial}{\partial x_1} S_{12}(x_1, x_2) \leq 0 \) and \( \frac{\partial}{\partial x_2} S_{12}(x_1, x_2) \leq 0 \)
4. \( \frac{\partial^2}{\partial x_2 \partial x_1} S_{12}(x_1, x_2) = f_{12}(x_1, x_2) \geq 0 \)

### 4.1.1 Bivariate Weibull-Gamma Distributions

We construct a bivariate Weibull survival function based on Model (4.1) and based on univariate Weibull survival functions, i.e. \( S_i(x_i) = \exp(-\lambda_i x_i^{\beta_i}) \), where \( \lambda_i > 0 \) and \( \beta_i > 0 \), \( i = 1, 2 \). The hazard function is \( h_i(x_i) = \lambda_i \beta_i x_i^{\beta_i - 1} \), and the cumulative hazard function is \( H_i(x_i) = \lambda_i x_i^{\beta_i} \), \( i = 1, 2 \). We assume \( g(u) = 1 - (1 + u) \exp(-u) \), but it is straightforward to invoke any of the other \( g(\cdot) \) functions described in section 4.1.

Therefore, the bivariate Weibull survival function is

\[
S_{12}(x_1, x_2) = S_1(x_1)S_2(x_2) \left[ 1 + \alpha_{12} g_1(H_1(x_1))g_2(H_2(x_2)) \right] = \exp\left(-\lambda_1 x_1^{\beta_1} - \lambda_2 x_2^{\beta_2}\right) \left[ 1 + \alpha_{12} \left(1 - (1 + \lambda_1 x_1^{\beta_1}) \exp(-\lambda_1 x_1^{\beta_1})\right) \right] \times \left(1 - (1 + \lambda_2 x_2^{\beta_2}) \exp(-\lambda_2 x_2^{\beta_2})\right) \tag{4.3}
\]

Moreover, the bivariate Weibull density function is

\[
f_{12}(x_1, x_2) = f_1(x_1)f_2(x_2) \left\{ 1 + \alpha_{12} \left(g_1(H_1(x_1)) - g_1'(H_1(x_1))\right) \times \left(g_2(H_2(x_2)) - g_2'(H_2(x_2))\right) \right\}
\]

\[
\left( \beta_1 \lambda_1 x_1^{\beta_1 - 1} \right) \left( \beta_2 \lambda_2 x_2^{\beta_2 - 1} \right) \exp(-\lambda_1 x_1^{\beta_1} - \lambda_2 x_2^{\beta_2}) \times \left[ 1 + \alpha_{12} \left(1 - \exp(-\lambda_1 x_1^{\beta_1}) - 2\lambda_1 x_1^{\beta_1} \exp(-\lambda_1 x_1^{\beta_1})\right) \right] \times \left(1 - \exp(-\lambda_2 x_2^{\beta_2}) - 2\lambda_2 x_2^{\beta_2} \exp(-\lambda_2 x_2^{\beta_2})\right)
\]

An appealing property of this bivariate Weibull density function is that if \( \alpha_{12} = 0 \), then the bivariate Weibull density function factors into the product of two independent univariate Weibull density functions, i.e.,

\[
f_{12}(x_1, x_2) = f_1(x_1)f_2(x_2)
\]

\[
= (\beta_1 \lambda_1 x_1^{\beta_1 - 1}) \exp(-\lambda_1 x_1^{\beta_1})(\beta_2 \lambda_2 x_2^{\beta_2 - 1}) \exp(-\lambda_2 x_2^{\beta_2})
\]

Thus, the two random variables are statistically independent. Another appealing property is that \( S_{12}(x_1, 0) = S_1(x_1) \) and \( S_{12}(0, x_2) = S_2(x_2) \), indicating that the marginal
survival functions also are univariate Weibull survival functions. The cumulative distribution function for the Weibull distribution is $F_i(x_i) = 1 - \exp(-\lambda_i x_i^{\beta_i}), i = 1, 2$.

Therefore, 

$$E(X_i) = \left(\frac{1}{\lambda_i}\right)^{\frac{1}{\beta_i}} \Gamma \left(1 + \frac{1}{\beta_i}\right)$$

and

$$Var(X_i) = \left(\frac{1}{\lambda_i}\right)^{\frac{2}{\beta_i}} \left\{\Gamma \left(1 + \frac{2}{\beta_i}\right) - \Gamma^2 \left(1 + \frac{1}{\beta_i}\right)\right\}$$

where

$$\Gamma(\alpha) = \int_0^\infty \exp(-x)x^{\alpha-1} \, dx, \quad \alpha > 0, i = 1, 2.$$

Next, we need to find the $Cov(X_1, X_2) = E(X_1X_2) - E(X_1)E(X_2)$. We first will find $E(X_1X_2)$.

$$E(X_1X_2) = \int_0^\infty \int_0^\infty x_1x_2f_{12}(x_1, x_2) \, dx_2 \, dx_1$$

$$= \left[\left(\frac{1}{\lambda_1}\right)^{\frac{1}{\beta_1}} \Gamma \left(1 + \frac{1}{\beta_1}\right) \left(\frac{1}{\lambda_2}\right)^{\frac{1}{\beta_2}} \Gamma \left(1 + \frac{1}{\beta_2}\right)\right] +$$

$$\left[\alpha_{12} \left(\frac{1}{\lambda_1}\right)^{\frac{1}{\beta_1}} \Gamma \left(1 + \frac{1}{\beta_1}\right) - \frac{1}{2} \left(\frac{1}{2\lambda_1}\right)^{\frac{1}{\beta_1}} \Gamma \left(1 + \frac{1}{\beta_1}\right) - \frac{1}{2} \left(\frac{1}{2\lambda_1}\right)^{\frac{1}{\beta_2}} \Gamma \left(1 + \frac{1}{\beta_2}\right)\right]$$

$$\times \left[\left(\frac{1}{\lambda_2}\right)^{\frac{1}{\beta_2}} \Gamma \left(1 + \frac{1}{\beta_2}\right) - \frac{1}{2} \left(\frac{1}{2\lambda_2}\right)^{\frac{1}{\beta_2}} \Gamma \left(1 + \frac{1}{\beta_2}\right) - \frac{1}{2} \left(\frac{1}{2\lambda_2}\right)^{\frac{1}{\beta_2}} \Gamma \left(1 + \frac{1}{\beta_2}\right)\right]$$

Therefore, we have

$$Cov(X_1, X_2) = E(X_1X_2) - E(X_1)E(X_2)$$

$$= \left[\alpha_{12} \left(\frac{1}{\lambda_1}\right)^{\frac{1}{\beta_1}} \Gamma \left(1 + \frac{1}{\beta_1}\right) - \frac{1}{2} \left(\frac{1}{2\lambda_1}\right)^{\frac{1}{\beta_1}} \Gamma \left(1 + \frac{1}{\beta_1}\right) - \frac{1}{2} \left(\frac{1}{2\lambda_1}\right)^{\frac{1}{\beta_2}} \Gamma \left(1 + \frac{1}{\beta_2}\right)\right]$$

$$\times \left[\left(\frac{1}{\lambda_2}\right)^{\frac{1}{\beta_2}} \Gamma \left(1 + \frac{1}{\beta_2}\right) - \frac{1}{2} \left(\frac{1}{2\lambda_2}\right)^{\frac{1}{\beta_2}} \Gamma \left(1 + \frac{1}{\beta_2}\right) - \frac{1}{2} \left(\frac{1}{2\lambda_2}\right)^{\frac{1}{\beta_2}} \Gamma \left(1 + \frac{1}{\beta_2}\right)\right]$$

Also,

$$Corr(X_1, X_2) = \frac{Cov(X_1, X_2)}{\sqrt{Var(X_1)Var(X_2)}}$$

(4.5)
where
\[ \text{Var}(X_i) = \left( \frac{1}{\lambda_i} \right)^2 \frac{2}{\beta_i} \left\{ \Gamma \left( 1 + \frac{2}{\beta_i} \right) - \Gamma^2 \left( 1 + \frac{1}{\beta_i} \right) \right\} \quad \text{for } i = 1, 2 \]

Notice that in the case of marginal exponential distributions (\( \beta_i = 1, \lambda_i = 1 \) for \( i = 1, 2 \)), \( \text{Corr}(X_1, X_2) = \frac{\alpha_{12}}{16} \). Because \(-1 \leq \alpha_{12} < +1\), this bivariate distribution experiences a similar drawback as the Farlie-Gumbel-Morgenstern family of bivariate distributions, i.e., it cannot accommodate high levels of correlation.

In case \( \alpha_{12} = 0 \), we find that
\[ E(X_1 X_2) = E(X_1) E(X_2) \]

Therefore, if \( \alpha_{12} = 0 \), we have
\[ \text{Cov}(X_1, X_2) = E(X_1 X_2) - E(X_1) E(X_2) = E(X_1) E(X_2) - E(X_1) E(X_2) = 0 \]

Hence, \( \text{Cov}(X_1, X_2) = \text{Corr}(X_1, X_2) = 0 \)

### 4.1.2 Bivariate Farlie-Gumbel-Morgenstern-Weibull Distributions

Let
\[ \mathbf{X} = \begin{bmatrix} X_1 & X_2 \end{bmatrix}^T \]
denote a random vector such that each \( X_i \) is a random variable, \( i = 1, 2 \). Let

- \( f_i(x_i) \) denote the density function
- \( F_i(x_i) = \int_0^{x_i} f_i(u) \, du \) denote the cumulative distribution function
- \( S_i(x_i) = 1 - F_i(x_i) = \int_{x_i}^{\infty} f_i(u) \, du \) denote the survival function

for the random variable \( X_i, i = 1, 2 \). All of these functions are nonnegative-valued, \( S_i(x_i) \) is a decreasing function of \( x_i \), and \( F_i(x_i) \) is an increasing function of \( x_i, i = 1, 2 \).
The Farlie-Gumbel-Morgenstern construction of a bivariate cumulative distribution function is

\[ F_{12}(x_1, x_2) = F_1(x_1) F_2(x_2) \left[ 1 + \alpha_{12} S_1(x_1) S_2(x_2) \right] \] (4.6)

where \( \alpha_{12} \) is an association or a correlation parameter, \(-1 \leq \alpha_{12} \leq +1\). Notice that \( \alpha_{12} = 0 \) yields the independence of \( X_1 \) and \( X_2 \). Also, notice that \( F_{12}(x_1, x_2) \to F_1(x_1) \) as \( x_2 \to \infty \), and \( F_{12}(x_1, x_2) \to F_2(x_2) \) as \( x_1 \to \infty \). Thus, the marginal cumulative distribution functions coincide with the original univariate cumulative distribution functions.

The bivariate density function is

\[ f_{12}(x_1, x_2) = (-1)^2 \frac{\partial^2}{\partial x_2 \partial x_1} F_{12}(x_1, x_2) \]

\[ = \frac{\partial^2}{\partial x_2 \partial x_1} F_1(x_1) F_2(x_2) \left[ 1 + \alpha_{12} S_1(x_1) S_2(x_2) \right] \]

\[ = \frac{\partial}{\partial x_1} F_1(x_1) \left[ F_2(x_2) \alpha_{12} S_1(x_1) (-f_2(x_2)) + \left( 1 + \alpha_{12} S_1(x_1) S_2(x_2) \right) f_2(x_2) \right] \]

\[ = -\rho_{12} F_1(x_1) f_2(x_2) \left( \frac{\partial}{\partial x_1} F_1(x_1) S_1(x_1) \right) \]

\[ + f_2(x_2) \left( \frac{\partial}{\partial x_1} F_1(x_1) (1 + \alpha_{12} S_1(x_1) S_2(x_2)) \right) \]

\[ = f_1(x_1) f_2(x_2) \left[ \left( 1 + \alpha_{12} \left( F_1(x_1) F_2(x_2) - S_1(x_1) F_2(x_2) - F_1(x_1) S_2(x_2) + S_1(x_1) S_2(x_2) \right) \right) \right] \]

\[ = f_1(x_1) f_2(x_2) \left[ 1 + \alpha_{12} \left( 1 - 2 F_1(x_1) \right) \left( 1 - 2 F_2(x_2) \right) \right] \]

Therefore, the bivariate Farlie-Gumbel-Morgenstern density function under the univariate cumulative Weibull distribution functions and the univariate Weibull density function is

\[ f_{12}(x_1, x_2) = (\beta_1 \lambda_1 x_1^{\beta_1 - 1}) \exp(-\lambda_1 x_1^{\beta_1}) (\beta_2 \lambda_2 x_2^{\beta_2 - 1}) \exp(-\lambda_2 x_2^{\beta_2}) \]

\[ \times \left[ 1 + \alpha_{12} \left( 1 - 2 \exp(-\lambda_1 x_1^{\beta_1}) \right) \left( 1 - 2 \exp(-\lambda_2 x_2^{\beta_2}) \right) \right] \]
An appealing property of this bivariate Weibull density function is that if $\alpha_{12} = 0$, then the bivariate Weibull density function factors into the product of two independent univariate Weibull density functions, i.e.,

$$f_{12}(x_1, x_2) = f_1(x_1)f_2(x_2) = (\beta_1 \lambda_1 x_1^{\beta_1-1}) \exp(-\lambda_1 x_1^{\beta_1})(\beta_2 \lambda_2 x_2^{\beta_2-1}) \exp(-\lambda_2 x_2^{\beta_2})$$

Thus, the two random variables are statistically independent.

Therefore,

$$E(X_i) = \left(\frac{1}{\lambda_i}\right)^{\frac{1}{\beta_i}} \Gamma\left(1 + \frac{1}{\beta_i}\right) \quad \text{and} \quad Var(X_i) = \left(\frac{1}{\lambda_i}\right)^{\frac{2}{\beta_i}} \left\{\Gamma\left(1 + \frac{2}{\beta_i}\right) - \Gamma^2\left(1 + \frac{1}{\beta_i}\right)\right\}$$

where

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} \exp(-x) \, dx, \quad \alpha > 0, \, i = 1, 2.$$

Next, we will find the covariance between $X$ and $Y$ for the bivariate Farlie-Gumbel-Morgenstern density function.

$$E(X_1X_2) = \int_0^\infty \int_0^\infty x_1 x_2 f_{12}(x_1, x_2) \, dx_1 \, dx_2$$

$$= \left[\int_0^\infty x_1 (\beta_1 \lambda_1 x_1^{\beta_1-1}) \exp(-\lambda_1 x_1^{\beta_1}) \int_0^\infty x_2 (\beta_2 \lambda_2 x_2^{\beta_2-1}) \exp(-\lambda_2 x_2^{\beta_2}) \, dx_2 \, dx_1\right]$$

$$+ \alpha_{12} \int_0^\infty x_1 (\beta_1 \lambda_1 x_1^{\beta_1-1}) \exp(-\lambda_1 x_1^{\beta_1}) \left(1 - 2 \exp(-\lambda_1 x_1^{\beta_1})\right) \left(1 - 2 \exp(-\lambda_2 x_2^{\beta_2})\right) \, dx_2 \, dx_1$$

$$= \left[\left(\frac{1}{\lambda_1}\right)^{\frac{1}{\beta_1}} \Gamma\left(1 + \frac{1}{\beta_1}\right) \left(\frac{1}{\lambda_2}\right)^{\frac{1}{\beta_2}} \Gamma\left(1 + \frac{1}{\beta_2}\right)\right]$$

$$+ \alpha_{12} \left(\frac{1}{\lambda_1}\right)^{\frac{1}{\beta_1}} \Gamma\left(1 + \frac{1}{\beta_1}\right) \left[\left(\frac{1}{2}\right)^{\frac{1}{\beta_1}} - 1\right] \left(\frac{1}{\lambda_2}\right)^{\frac{1}{\beta_2}} \Gamma\left(1 + \frac{1}{\beta_2}\right) \left[\left(\frac{1}{2}\right)^{\frac{1}{\beta_2}} - 1\right]$$

$$= \left(\frac{1}{\lambda_1}\right)^{\frac{1}{\beta_1}} \left(\frac{1}{\lambda_2}\right)^{\frac{1}{\beta_2}} \Gamma\left(1 + \frac{1}{\beta_1}\right) \Gamma\left(1 + \frac{1}{\beta_2}\right) \left[1 + \alpha_{12} \left(\frac{1}{2}\right)^{\frac{1}{\beta_1}} - 1\right] \left(\frac{1}{2}\right)^{\frac{1}{\beta_2}} - 1\right]$$
Therefore, we have

\[ \text{Cov}(X_1, X_2) = E(X_1X_2) - E(X_1)E(X_2) \]
\[ = \left( \frac{1}{\lambda_1} \right)^{\frac{1}{\beta_1}} \left( \frac{1}{\lambda_2} \right)^{\frac{1}{\beta_2}} \Gamma \left( 1 + \frac{1}{\beta_1} \right) \Gamma \left( 1 + \frac{1}{\beta_2} \right) \]
\[ \times \left( \alpha_{12} \left( \left( \frac{1}{2} \right)^{\frac{1}{\beta_1}} - 1 \right) \left( \left( \frac{1}{2} \right)^{\frac{1}{\beta_2}} - 1 \right) \right) \]

Also,

\[ \text{Corr}(X_1, X_2) = \frac{\text{Cov}(X_1, X_2)}{\sqrt{\text{Var}(X_1)\text{Var}(X_2)}} \] \hspace{1cm} (4.7)

where

\[ \text{Var}(X_i) = \left( \frac{1}{\lambda_i} \right)^{\frac{2}{\beta_i}} \left\{ \Gamma \left( 1 + \frac{2}{\beta_i} \right) - \Gamma^2 \left( 1 + \frac{1}{\beta_i} \right) \right\} \quad \text{for} \quad i = 1, 2 \]

Notice that in the case of marginal exponential distributions \( \beta_i = 1, \lambda_i = 1 \) for \( i = 1, 2 \), \( \text{Corr}(X_1, X_2) = \frac{\alpha_{12}}{4} \). Because \( -1 \leq \alpha_{12} < +1 \), the Farlie-Gumbel-Morgenstern bivariate distribution experiences the drawback that it cannot accommodate high levels of correlation.

In case \( \rho_{12} = 0 \), we find that

\[ E(X_1X_2) = E(X_1)E(X_2) \]

Therefore, if \( \rho_{12} = 0 \), we have

\[ \text{Cov}(X_1, X_2) = E(X_1X_2) - E(X_1)E(X_2) \]
\[ = E(X_1)E(X_2) - E(X_1)E(X_2) \]
\[ = 0 \]

Hence, \( \text{Cov}(X_1, X_2) = \text{Corr}(X_1, X_2) = 0 \).
4.1.3 Bivariate Piecewise Uniform-Weibull Distribution

Another approach to accommodating left-censored data is to assume a uniform distribution in the region below the LLD and a distribution, such as a Weibull, in the region above the LLD.

Let $X$ be a nonnegative-valued random variable such that it is a mixture of a uniform random variable and a Weibull random variable, defined as

$$X = I(0 \leq X \leq LLD) \cdot U + I(X > LLD) \cdot W$$

where $LLD$ represents the lower limit of detection (a known positive value) and $I(A)$ represents the indicator function that yields 1 or 0 depending on whether the event $A$ holds. Then the density function for $X$ is

$$f(x) = \theta \cdot I(0 \leq x \leq LLD)\left[\frac{1}{LLD}\right] + (1 - \theta) \cdot I(x > LLD)\left[\lambda \beta (x - LLD)^{\beta - 1} \exp(-\lambda(x - LLD)^{\beta})\right]$$

where $\theta = Pr(0 \leq X \leq LLD)$.

The cumulative distribution function (cdf) for $X$ is

$$F(x) = \theta \cdot I(0 \leq x \leq LLD)\left[\frac{x}{LLD}\right] + I(x > LLD)\left[\theta + (1 - \theta)[1 - \exp(-\lambda(x - LLD)^{\beta})]\right]$$

The moments of $X$ are derived as follows:

$$E(X) = \theta\left(\frac{LLD}{2}\right) + (1 - \theta)(E(W) + LLD)$$

where $W \sim \text{Weibull}(\lambda, \beta)$ and $E(W) = \left(\frac{1}{\lambda}\right)^{\frac{1}{\beta}} \Gamma\left(1 + \frac{1}{\beta}\right)$

$$E(X^2) = \theta\left(\frac{LLD^2}{3}\right) + (1 - \theta)[E(W^2) + 2 \cdot LLN \cdot E(W) + LLN^2]$$
where \( E(W^2) = \left( \frac{1}{\lambda} \right)^{\frac{2}{\beta}} \Gamma \left( 1 + \frac{2}{\beta} \right) \)

\[
Var(X) = E(X^2) - (E(X))^2
\]

In order to calculate the covariance in a bivariate Farlie-Gumbel-Morgenstern (FGM) distribution, \( E(X \cdot F(X)) \) is needed.

\[
E(X \cdot F(X)) = \int x \cdot f(x) \cdot F(x) dx
= \int I(0 \leq x \leq LLD) x f(X) F(X) dx + \int I(X > LLD) x f(X) F(X) dx
= \int I(0 \leq x \leq LLD) \theta^2 x^2 \left( \frac{1}{LLD^2} \right) dx
+ \int I(X > LLD) x (1 - \theta) f_W(x - LLD) \left[ \theta + (1 - \theta) F_W(x - LLD) \right] dx
\]

where \( f_W(x) \) denotes the Weibull density and \( F_W(x) \) denotes the Weibull cdf. Then

\[
E(X \cdot F(X)) = \theta^2 \left( \frac{LLD}{3} \right) + \theta (1 - \theta) \int I(X > LLD) x f_W(x - LLD) dx
+ (1 - \theta)^2 \int I(X > LLD) x \cdot f_W(x - LLD) F_W(x - LLD) \left[ \theta + (1 - \theta) F_W(x - LLD) \right] dx
= \theta^2 \left( \frac{LLD}{3} \right) + \theta (1 - \theta) (E(W) + LLD)
+ (1 - \theta)^2 \int I(X > LLD) x \cdot f_W(x - LLD) F_W(x - LLD) \left[ \theta + (1 - \theta) F_W(x - LLD) \right] dx
= \theta^2 \left( \frac{LLD}{3} \right) + (1 - \theta) (E(W) + LLD)
- \frac{(1 - \theta)^2}{2} (E(W^*) + LLD)
\]

where \( W^* \sim Weibull(2\lambda, \beta) \) and \( E(W^*) = \left( \frac{1}{2\lambda} \right)^{\frac{1}{\beta}} \Gamma \left( 1 + \frac{1}{\beta} \right) \)

For the bivariate FGM distribution, the bivariate cdf is

\[
F_{12}(x_1, x_2) = F_1(x_1) F_2(x_2) \left[ 1 + \alpha_{12}(1 - F_1(x_1))(1 - F_2(x_2)) \right]
\]
and the bivariate density function is

\[ f_{12}(x_1, x_2) = f_1(x_1) f_2(x_2) \left[ 1 + \alpha_{12} \left( 1 - 2F_1(x_1) \right) \left( 1 - 2F_2(x_2) \right) \right] \]

where \(-1 < \alpha_{12} < +1\).

Then

\[
Cov(X_1, X_2) = E(X_1X_2) - E(X_1)E(X_2)
\]

\[
= \alpha_{12} \int \int x_1 x_2 \cdot f_1(x_1) f_2(x_2) \left( 1 - 2F_1(x_1) \right) \left( 1 - 2F_2(x_2) \right) dx_1dx_2
\]

\[
= \alpha_{12} \left( \int x_1 f_1(x_1) \left( 1 - 2F_1(x_1) \right) dx_1 \right) \left( \int x_2 f_2(x_2) \left( 1 - 2F_2(x_2) \right) dx_2 \right)
\]

\[
= \alpha_{12} \left( \int x_1 f_1(x_1) dx_1 - 2 \int x_1 f_1(x_1) F_1(x_1) dx_1 \right)
\]

\[
\times \left( \int x_2 f_2(x_2) dx_2 - 2 \int x_2 f_2(x_2) F_2(x_2) dx_2 \right)
\]

\[
= \alpha_{12} \left[ \theta_X \left( \frac{LLD_X}{2} \right) - (1 - \theta_X) \left( \left( \frac{1}{\lambda_X} \right)^{\frac{1}{\beta_X}} \Gamma \left( 1 + \frac{1}{\beta_X} \right) + LLD_X \right) \right]
\]

\[
- 2 \left( \theta_X^2 \left( \frac{LLD_X}{3} \right) - (1 - \theta_X)^2 \left( \left( \frac{1}{2\lambda_X} \right)^{\frac{1}{\beta_X}} \Gamma \left( 1 + \frac{1}{\beta_X} \right) + LLD_X \right) \right)
\]

\[
\times \left[ \theta_Y \left( \frac{LLD_Y}{2} \right) - (1 - \theta_Y) \left( \left( \frac{1}{\lambda_Y} \right)^{\frac{1}{\beta_Y}} \Gamma \left( 1 + \frac{1}{\beta_Y} \right) + LLD_Y \right) \right]
\]

\[
- 2 \left( \theta_Y^2 \left( \frac{LLD_Y}{3} \right) - (1 - \theta_Y)^2 \left( \left( \frac{1}{2\lambda_Y} \right)^{\frac{1}{\beta_Y}} \Gamma \left( 1 + \frac{1}{\beta_Y} \right) + LLD_Y \right) \right)
\]

4.2 Multivariate Survival Functions

Model (4.1) can be extended to the situation with \( p \) random variables as
4.2.1 Multivariate Weibull Density Function

From model (4.6), we consider the special case of the cumulative hazard function

\[ \lambda \]

and based on univariate Weibull survival functions, i.e.

\[ S_{12} > \ldots p \]

When we assume the multivariate Weibull survival function based on Model (4.8)

Expanding this leads to

Correspondingly, the multivariate Weibull survival function is

When we assume the multivariate Weibull survival function based on Model (4.8) and based on univariate Weibull survival functions, i.e. \( S_i(x_i) = \exp(-\lambda_i x_i^{\beta_i}) \), where \( \lambda_i > 0 \) and \( \beta_i > 0 \), \( i = 1, 2, \ldots, p \). The hazard function is \( h_i(x_i) = \lambda_i \beta_i x_i^{\beta_i-1} \), and the cumulative hazard function is \( H_i(x_i) = \lambda_i x_i^{\beta_i}, i = 1, 2, \ldots, p \).

4.2.1 Multivariate Weibull Density Function

From model (4.6), we consider the special case of \( p = 3 \), and drop the arguments \( x_1, x_2, x_3 \) for notational convenience. Then the trivariate survival function is

Expanding this leads to

\[
S_{123} = S_1 S_2 S_3 \left( 1 + \alpha_{12} g_1 g_2 \right) \left( 1 + \alpha_{13} g_1 g_3 \right) \left( 1 + \alpha_{23} g_2 g_3 \right)
\]

\[
S_{123} = S_1 S_2 S_3 \left( 1 + \alpha_{12} g_1 g_2 + \alpha_{13} g_1 g_3 + \alpha_{23} g_2 g_3 + \alpha_{12} \alpha_{13} g_1^2 g_2 g_3 + \alpha_{12} \alpha_{23} g_1 g_2^2 g_3 + \alpha_{12} \alpha_{13} \alpha_{23} g_1^2 g_2^2 g_3 \right)
\]
To identify the trivariate density function, determine \( \frac{\partial^2}{\partial x_1 \partial x_2} S_{123}(x_1, x_2, x_3) \). The first partial derivative of the trivariate survival function with respect to \( x_1 \) is

\[
\frac{\partial}{\partial x_1} S_{123} = S_1 S_2 S_3 \left( \alpha_{12} g_2 g_1' h_1 + \alpha_{13} g_3 g_1' h_1 \\
+ \alpha_{12} \alpha_{13} g_2 g_3 g_1' h_1 + \alpha_{12} \alpha_{23} g_2^2 g_3 g_1' h_1 + \alpha_{13} \alpha_{23} g_2^2 g_1 h_1 \\
+ \alpha_{12} \alpha_{13} \alpha_{23} g_2^2 g_3^2 g_1' h_1 \right) \\
- f_1 S_2 S_3 \left( 1 + \alpha_{12} g_1 g_2 + \alpha_{13} g_1 g_3 + \alpha_{23} g_2 g_3 \\
+ \alpha_{12} \alpha_{13} g_2 g_3 + \alpha_{12} \alpha_{23} g_1 g_2 g_3 + \alpha_{13} \alpha_{23} g_1 g_2 g_3 \\
+ \alpha_{12} \alpha_{13} \alpha_{23} g_1^2 g_2 g_3 \right)
\]

which reduces to

\[
\frac{\partial}{\partial x_1} S_{123} = f_1 S_2 S_3 \left( -1 + \alpha_{12} (g_1' - g_1) g_2 + \alpha_{13} (g_1' - g_1) g_3 + \alpha_{23} g_2 g_3 \\
+ \alpha_{12} \alpha_{13} (2g_1 g_1' - g_1^2) g_2 g_3 + \alpha_{12} \alpha_{23} (g_1' - g_1) g_2 g_3 \\
+ \alpha_{13} \alpha_{23} (g_1' - g_1) g_2 g_3 + \alpha_{12} \alpha_{13} \alpha_{23} (2g_1 g_1' - g_1^2) g_2 g_3 \right)
\]

The second partial derivative of the trivariate survival function with respect to \( x_1 \) and \( x_2 \) is

\[
\frac{\partial^2}{\partial x_1 \partial x_2} S_{123} = f_1 S_2 S_3 \left( \alpha_{12} (g_1' - g_1) g_2 h_2 - \alpha_{23} g_3 g_2' h_2 + \alpha_{12} \alpha_{13} (2g_1' - g_1) g_1 g_3 g_2 h_2 \\
+ \alpha_{12} \alpha_{23} (g_1' - g_1) g_3 g_2 g_2' h_2 + \alpha_{13} \alpha_{23} (g_1' - g_1) g_3^2 g_2 h_2 \\
+ \alpha_{12} \alpha_{13} \alpha_{23} (2g_1' - g_1) g_1 g_3^2 g_2 g_2' h_2 \right) \\
+ f_1 f_2 S_3 \left( 1 + \alpha_{12} (g_1 - g_1') g_2 + \alpha_{13} (g_1 - g_1') g_3 + \alpha_{23} g_2 g_3 \\
+ \alpha_{12} \alpha_{13} (g_1 - 2g_1') g_1 g_2 g_3 + \alpha_{12} \alpha_{23} (g_1 - g_1') g_2^2 g_3 \\
+ \alpha_{13} \alpha_{23} (g_1 - g_1') g_2 g_3 + \alpha_{12} \alpha_{13} \alpha_{23} (g_1 - 2g_1') g_1 g_2 g_3 \right)
\]
which reduces to
\[
\frac{\partial^2}{\partial x_1 \partial x_2} S_{123} = f_1 f_2 S_3 \left( 1 + \alpha_{12}(g_1 - g_1')(g_2 - g_2') + \alpha_{13}(g_1 - g_1')(g_3 - g_3') + \alpha_{23}(g_2 - g_2')(g_3 - g_3') \\
+ \alpha_{12}\alpha_{13}(g_1 - 2g_1')(g_2 - g_2')g_1g_3 + \alpha_{12}\alpha_{23}(g_1 - g_1')(g_2 - 2g_2')g_2g_3 \\
+ \alpha_{13}\alpha_{23}(g_1 - g_1')(g_2 - g_2')g_3^2 + \alpha_{12}\alpha_{13}\alpha_{23}(g_1 - 2g_1')(g_2 - 2g_2')g_1g_2g_3^2 \right)
\]

The third partial derivative of the trivariate survival function with respect to \(x_1, x_2,\) and \(x_3\) is
\[
\frac{\partial^3}{\partial x_1 \partial x_2 \partial x_3} S_{123} = f_1 f_2 S_3 \left( \alpha_{23}(g_2 - g_2')g_3^2 h_3 + \alpha_{13}(g_1 - g_1')g_3^2 h_3 \\
+ \alpha_{12}\alpha_{13}(g_1 - 2g_1')(g_2 - g_2')g_1g_3^2 h_3 \\
+ \alpha_{12}\alpha_{23}(g_1 - g_1')(g_2 - 2g_2')g_2g_3^2 h_3 \\
+ \alpha_{13}\alpha_{23}(g_1 - g_1')g_3^2 h_3 \\
+ \alpha_{12}\alpha_{13}\alpha_{23}(g_1 - 2g_1')(g_2 - 2g_2')g_1g_2g_3^2 \right)
\]
\[
+ f_1 f_2 f_3 \left( 1 + \alpha_{12}(g_1 - g_1')(g_2 - g_2') + \alpha_{13}(g_1 - g_1')(g_3 - g_3') + \alpha_{23}(g_2 - g_2')(g_3 - g_3') \\
+ \alpha_{12}\alpha_{13}(g_1 - 2g_1')(g_2 - g_2')g_1g_3 + \alpha_{12}\alpha_{23}(g_1 - g_1')(g_2 - 2g_2')g_2g_3 \\
+ \alpha_{12}\alpha_{23}(g_1 - g_1')(g_2 - g_2')g_3^2 + \alpha_{12}\alpha_{13}\alpha_{23}(g_1 - 2g_1')(g_2 - 2g_2')g_1g_2g_3^2 \right)
\]

Therefore, the trivariate density function is
\[
f_{123} = (-1)^3 \frac{\partial^3}{\partial x_1 \partial x_2 \partial x_3} S_{123} \\
= f_1 f_2 f_3 \left( 1 + \alpha_{12}(g_1 - g_1')(g_2 - g_2') + \alpha_{13}(g_1 - g_1')(g_3 - g_3') + \alpha_{23}(g_2 - g_2')(g_3 - g_3') \\
+ \alpha_{12}\alpha_{13}(g_1 - 2g_1')(g_2 - g_2')(g_3 - g_3')g_1 \\
+ \alpha_{12}\alpha_{23}(g_1 - g_1')(g_2 - 2g_2')(g_3 - g_3')g_2 \\
+ \alpha_{13}\alpha_{23}(g_1 - g_1')(g_2 - g_2')(g_3 - 2g_3')g_3 \\
+ \alpha_{12}\alpha_{13}\alpha_{23}(g_1 - 2g_1')(g_2 - 2g_2')(g_3 - 2g_3')g_1g_2g_3 \right)
\]
To confirm that this indeed is a density function, let \( O(-1, +1) \) denote any function that is bounded between \(-1\) and \(+1\). Notice that

\[
(g_i - g'_i) = O(-1, +1) \quad \text{and} \quad (g_i - g'_i)(g_i - 2g'_i) = O(-1, +1), \quad i = 1, 2, 3.
\]

Then the trivariate density function can be expressed as

\[
f_{123} = f_1f_2f_3\left(1 + \alpha_{12}O(-1, +1)\right)\left(1 + \alpha_{13}O(-1, +1)\right)\left(1 + \alpha_{23}O(-1, +1)\right)
\]

which is a nonnegative-valued function because each of \( f_1, f_2, \) and \( f_3 \) is nonnegative, \( 0 \leq g_i, g'_i \leq 1 \) for \( i = 1, 2, 3 \), and \(-1 \leq \alpha_{12} \leq +1, -1 \leq \alpha_{13} \leq +1, \) and \(-1 \leq \alpha_{23} \leq +1\).

4.2.2 Multivariate Farlie-Gumbel-Morgenstern Density Function

Let

\[
X = \begin{bmatrix} X_1 & X_2 & \ldots & X_p \end{bmatrix}^T
\]

denote a random vector such that each \( X_i \) is a random variable, \( i = 1, 2, \ldots, p \). Construct a multivariate Farlie-Gumbel-Morgenstern cumulative distribution function as

\[
F_{12\ldots p}(x_1, x_2, \ldots, x_p) = F_1(x_1)F_2(x_2)\ldots F_p(x_p)\left(1 + \alpha_{12}S_1(x_1)S_2(x_2)\right)\times \left(1 + \alpha_{13}S_1(x_1)S_3(x_3)\right)\times \ldots \times \left(1 + \alpha_{p-1,p}S_{p-1}(x_{p-1})S_p(x_p)\right)
\]

where each \( \alpha_{ij} \) is a correlation parameter, \(-1 \leq \alpha_{ij} \leq +1, i = 1, 2, \ldots, p - 1 \) and \( j = i + 1, i + 2, \ldots, p \).

Consider the special case of \( p = 3 \), and drop the arguments \( x_1, x_2, x_3 \) for notational convenience. Then the trivariate cumulative distribution function is

\[
F_{123} = F_1F_2F_3\left(1 + \alpha_{12}S_1S_2\right)\left(1 + \alpha_{13}S_1S_3\right)\left(1 + \alpha_{23}S_2S_3\right)
\]
Expanding this leads to

\[
F_{123} = F_1 F_2 F_3 \left( 1 + \alpha_{12} S_1 S_2 + \alpha_{13} S_1 S_3 + \alpha_{23} S_2 S_3 \\
+ \alpha_{12} \alpha_{13} S_1^2 S_2 S_3 + \alpha_{12} \alpha_{23} S_1 S_2^2 S_3 + \alpha_{13} \alpha_{23} S_1 S_2 S_3^2 + \alpha_{12} \alpha_{13} \alpha_{23} S_1^2 S_2 S_3^2 \right)
\]

To identify the trivariate density function, determine \( \frac{\partial}{\partial x_1} F_{123}(x_1, x_2, x_3) \). The first partial derivative of the trivariate cumulative distribution function with respect to \( x_1 \) is

\[
\frac{\partial}{\partial x_1} F_{123} = f_1 F_2 F_3 \left( 1 + \alpha_{12} S_1 S_2 + \alpha_{13} S_1 S_3 + \alpha_{23} S_2 S_3 \\
+ \alpha_{12} \alpha_{13} S_1^2 S_2 S_3 + \alpha_{12} \alpha_{23} S_1 S_2^2 S_3 + \alpha_{13} \alpha_{23} S_1 S_2 S_3^2 + \alpha_{12} \alpha_{13} \alpha_{23} S_1^2 S_2 S_3^2 \right)
\]

\[
- f_1 F_2 F_3 \left( \alpha_{12} F_1 S_2 + \alpha_{13} F_1 S_3 + 2 \alpha_{12} \alpha_{13} F_1 S_1 S_2 S_3 \\
+ \alpha_{12} \alpha_{23} F_1 S_2^2 S_3 + \alpha_{13} \alpha_{23} F_1 S_2 S_3^2 + 2 \alpha_{12} \alpha_{13} \alpha_{23} F_1 S_1 S_2 S_3^2 \right)
\]

which reduces to

\[
\frac{\partial}{\partial x_1} F_{123} = f_1 F_2 F_3 \left( 1 + \alpha_{12}(1 - 2F_1)S_2 + \alpha_{13}(1 - 2F_1)S_3 + \alpha_{23} S_2 S_3 \\
+ \alpha_{12} \alpha_{13}(1 - F_1)(1 - 3F_1)S_2^2 S_3 + \alpha_{12} \alpha_{23}(1 - 2F_1)S_2^2 S_3 \\
+ \alpha_{13} \alpha_{23}(1 - 2F_1)S_2 S_3^2 + \alpha_{12} \alpha_{13} \alpha_{23}(1 - F_1)(1 - 3F_1)S_2^2 S_3^2 \right)
\]

The second partial derivative of the trivariate cumulative distribution function with respect to \( x_1 \) and \( x_2 \) is
\[
\frac{\partial^2}{\partial x_1 \partial x_2} F_{123} = f_1 f_2 F_3 \left( 1 + \alpha_{12}(1 - 2F_1)S_2 + \alpha_{13}(1 - 2F_1)S_3 + \alpha_{23}S_2S_3 \right)
\]
\[
+ \alpha_{12}\alpha_{13}(1 - F_1)(1 - 3F_1)S_2S_3 + \alpha_{12}\alpha_{23}(1 - 2F_1)S_2^2S_3
\]
\[
+ \alpha_{13}\alpha_{23}(1 - 2F_1)S_2S_3^2 + \alpha_{12}\alpha_{13}\alpha_{23}(1 - F_1)(1 - 3F_1)S_2^2S_3^2
\]
\[
- f_1 f_2 F_3 \left( \alpha_{12}(1 - 2F_1)F_2 + \alpha_{23}F_2S_3 + \alpha_{12}\alpha_{13}(1 - F_1)(1 - 3F_1)F_2S_3 
\right)
\]
\[
+ 2\alpha_{12}\alpha_{23}(1 - 2F_1)F_2S_2S_3 + \alpha_{13}\alpha_{23}(1 - 2F_1)F_2S_3^2 
\]
\[
+ 2\alpha_{12}\alpha_{13}\alpha_{23}(1 - F_1)(1 - 3F_1)F_2S_2S_3^2
\]

which reduces to

\[
\frac{\partial^2}{\partial x_1 \partial x_2} F_{123} = f_1 f_2 F_3 \left( 1 + \alpha_{12}(1 - 2F_1)(1 - 2F_2) + \alpha_{13}(1 - 2F_1)S_3 + \alpha_{23}(1 - 2F_2)S_3 \right)
\]
\[
+ \alpha_{12}\alpha_{13}(1 - F_1)(1 - 3F_1)(1 - 2F_2)S_3 + \alpha_{12}\alpha_{23}(1 - 2F_1)(1 - F_2)(1 - 3F_2)S_3 
\]
\[
+ \alpha_{13}\alpha_{23}(1 - 2F_1)(1 - 2F_2)S_3^2 
\]
\[
+ \alpha_{12}\alpha_{13}\alpha_{23}(1 - F_1)(1 - 3F_1)(1 - F_2)(1 - 3F_2)S_3^2 
\]

The third partial derivative of the trivariate cumulative distribution function with respect to \( x_1, x_2, \) and \( x_3 \) is
\[
\frac{\partial^3}{\partial x_1 \partial x_2 \partial x_3} F_{123} = f_1 f_2 f_3 \left( 1 + \alpha_{12}(1 - 2F_1)(1 - 2F_2) + \alpha_{13}(1 - 2F_1)S_3 + \alpha_{23}(1 - 2F_2)S_3 \\
+ \alpha_{12} \alpha_{13}(1 - F_1)(1 - 3F_1)(1 - 2F_2)S_3 \\
+ \alpha_{12} \alpha_{23}(1 - 2F_1)(1 - F_2)(1 - 3F_2)S_3 + \alpha_{13} \alpha_{23}(1 - 2F_1)(1 - 2F_2)S_3^2 \\
+ \alpha_{12} \alpha_{13} \alpha_{23}(1 - F_1)(1 - 3F_1)(1 - F_2)(1 - 2F_2)S_3^2 \right) \\
- f_1 f_2 f_3 \left( \alpha_{13}(1 - 2F_1)F_3 + \alpha_{23}(1 - 2F_2)F_3 \\
+ \alpha_{12} \alpha_{13}(1 - F_1)(1 - 3F_1)(1 - 2F_2)F_3 \\
+ \alpha_{12} \alpha_{23}(1 - 2F_1)(1 - F_2)(1 - 3F_2)F_3 + 2\alpha_{13} \alpha_{23}(1 - 2F_1)(1 - 2F_2)F_3S_3 \\
+ 2\alpha_{12} \alpha_{13} \alpha_{23}(1 - F_1)(1 - 3F_1)(1 - F_2)(1 - 3F_2)F_3S_3 \right)
\]

which reduces to

\[
\frac{\partial^3}{\partial x_1 \partial x_2 \partial x_3} F_{123} = f_1 f_2 f_3 \left( 1 + \alpha_{12}(1 - 2F_1)(1 - 2F_2) \\
+ \alpha_{13}(1 - 2F_1)(1 - 2F_3) + \alpha_{23}(1 - 2F_2)(1 - 2F_3) \\
+ \alpha_{12} \alpha_{13}(1 - F_1)(1 - 3F_1)(1 - 2F_2)(1 - 2F_3) \\
+ \alpha_{12} \alpha_{23}(1 - 2F_1)(1 - F_2)(1 - 3F_2)(1 - 2F_3) \\
+ \alpha_{13} \alpha_{23}(1 - 2F_1)(1 - 2F_2)(1 - F_3)(1 - 3F_3) \\
+ \alpha_{12} \alpha_{13} \alpha_{23}(1 - F_1)(1 - 3F_1)(1 - F_2)(1 - 3F_2)(1 - F_3)(1 - 3F_3) \right)
\]
Therefore, the trivariate density function is

\[ f_{123} = f_1 f_2 f_3 \left( 1 + \alpha_{12}(1 - 2F_1)(1 - 2F_2) \right. \]
\[ + \alpha_{13}(1 - 2F_1)(1 - 2F_3) + \alpha_{23}(1 - 2F_2)(1 - 2F_3) \]
\[ + \alpha_{12}\alpha_{13}(1 - F_1)(1 - 3F_1)(1 - 2F_2)(1 - 2F_3) \]
\[ + \alpha_{12}\alpha_{23}(1 - 2F_1)(1 - F_2)(1 - 3F_2)(1 - 2F_3) \]
\[ + \alpha_{13}\alpha_{23}(1 - 2F_2)(1 - F_3)(1 - 3F_2)(1 - 2F_3) \]
\[ + \alpha_{12}\alpha_{13}\alpha_{23}(1 - F_1)(1 - 3F_1)(1 - F_2)(1 - 3F_2)(1 - F_3)(1 - 3F_3) \]

To confirm that this indeed is a density function, let \( O(-1,+1) \) denote any function that is bounded between \(-1\) and \(+1\). Notice that

\[ (1 - 2F_i) = O(-1,+1) \quad \text{and} \quad (1 - F_i)(1 - 3F_i) = O(-1,+1), \ i = 1, 2, 3. \]

Then the trivariate density function can be expressed as

\[ f_{123} = f_1 f_2 f_3 \left( 1 + \alpha_{12}O(-1,+1) \right) \left( 1 + \alpha_{13}O(-1,+1) \right) \left( 1 + \alpha_{23}O(-1,+1) \right) \]

which is a nonnegative-valued function because each of \( f_1, f_2, \) and \( f_3 \) is nonnegative and \(-1 \leq \alpha_{12} \leq +1, -1 \leq \alpha_{13} \leq +1, \) and \(-1 \leq \alpha_{23} \leq +1.\)
Chapter 5  
Simulation Studies

We investigate how the maximum likelihood (ML) approach performs under four bivariate distributions (1) Bivariate Normal distribution (2) Bivariate Weibull-Gamma distribution (3) Bivariate FGM-Weibull Distribution, and (4) Piecewise Uniform-Weibull distribution in a computer simulation study. In all simulations described, we generate the paired data represented by the variables $X_1$ and $X_2$ with a sample size of 100, 50, 25 for each of 1000 data sets and left-censoring rates of (25% for $X_1$, 25% for $X_2$) or (40% for $X_1$, 25% for $X_2$) or (60% for $X_1$, 25% for $X_2$). The selected values of the LLDs in the simulation study are determined by the censoring rates. All calculations are performed using SAS 9.3 statistical software. In section 5.1, the paired samples were generated from the bivariate lognormal distribution. In section 5.2, the paired samples were generated from the bivariate FGM-Weibull distribution.

The purpose of this simulation study is to examine the robustness and performance of all four bivariate distributions when the data sets are generated by different distributions from the model used in the maximum likelihood (ML) approach. For example, if the data sets are generated from the Bivariate lognormal distribution, then we want to determine from the simulation study whether the Bivariate Weibull-Gamma ML model estimation, the Bivariate FGM-Weibull ML model estimation, and the Piecewise Uniform-Weibull ML model estimation perform well, i.e., are they robust to the underlying distributional assumptions.
5.1 Data Set Generated by a Bivariate Lognormal Distribution

In this section, the paired samples were generated from the bivariate lognormal distribution using SAS/IML software for each of the following four distributions:

1. Maximum likelihood method based on bivariate normal distribution
2. Maximum likelihood method based on bivariate Weibull - Gamma distribution
3. Maximum likelihood method based on bivariate FGM - Weibull distribution
4. Maximum likelihood method based on Piecewise Uniform - Weibull distribution

The scenarios considered here for our computer simulation studies are similar to those considered by Barnhart et al. [5]. We generate the paired data represented by the variables $X_1$ and $X_2$ with a sample size of 100, 50, 25 for each of 1000 data sets using one of the following six combinations of parameter settings for the means, standard deviations, and correlation coefficient: $\mu_{x_1} = 0, \mu_{x_2} = 0.2, \sigma_{x_1} = 0.8, \sigma_{x_2} = 1, \rho = 0.25, 0.50, 0.75$, and left-censoring rates of (25% for $X_1$, 25% for $X_2$) or (40% for $X_1$, 25% for $X_2$) or (60% for $X_1$, 25% for $X_2$).

To assess the level of agreement of a biomarker measured under two different conditions, we use Lin’s concordance correlation coefficient (CCC) index [3], expressed as

$$CCC = \rho_c = 1 - \frac{E[(X_1 - X_2)^2]}{E[(X_1 - X_2)^2|X_1, X_2 \text{are independent}] + 2COV(X_1, X_2)}$$

(5.1)

The maximum likelihood method (ML method) under a bivariate distribution for estimating the CCC in the presence of left-censored data can be obtained by inserting the ML estimates for $COV(X_1, X_2), VAR(X_1), VAR(X_2), E(X_1), E(X_2)$ into equation (5.1). To obtain the parameter estimates, we derive the likelihood function according to the lower limit of detection of each random variable $X_1$ and $X_2$.

Let $X_1$ and $X_2$ be continuous random variables representing the two assay readings on the same subject based on two different techniques. Define $LLD_{x_1}$ and $LLD_{x_2}$.
as the left-censored variables corresponding to $X_1$ and $X_2$. Let $(x_{1i}, x_{2i}), i = 1, ..., N$ be a random sample from random variables $(X_{1i}, X_{2i})$. The maximum likelihood estimates can be obtained by maximizing the observed data likelihood, based on $N$ pairs of $(X_1, X_2)$.

**Method 1: Maximum likelihood based on the bivariate lognormal distribution**

The maximum likelihood estimates can be obtained by maximizing the observed data likelihood, based on $N$ pairs of $(X_1, X_2)$. All estimated CCCs ($\hat{\rho}_c$) were obtained by maximizing the likelihood function with respect to each of the following four scenarios:

**Case 1:** $x_1 \geq LLD_{x_1}$ and $x_2 \geq LLD_{x_2}$

$$BVN \left( \frac{x_1 - \mu_{x_1}}{\sigma_{x_1}}, \frac{x_2 - \mu_{x_2}}{\sigma_{x_2}}, \sigma_{x_1x_2} \right)$$

**Case 2:** $x_1 < LLD_{x_1}$ and $x_2 \geq LLD_{x_2}$

$$\Phi \left( \frac{LLD_{x_1} - \mu_{x_1|x_2}}{\sigma_{x_1|x_2}} \right) \phi \left( \frac{x_2 - \mu_{x_2}}{\sigma_{x_2}} \right)$$

**Case 3:** $x_1 \geq LLD_{x_1}$ and $x_2 < LLD_{x_2}$

$$\phi \left( \frac{x_1 - \mu_{x_1}}{\sigma_{x_1}} \right) \Phi \left( \frac{LLD_{x_2} - \mu_{x_2|x_1}}{\sigma_{x_2|x_1}} \right)$$

**Case 4:** $x_1 < LLD_{x_1}$ and $x_2 < LLD_{x_2}$

$$BVN \left( \frac{LLD_{x_1} - \mu_{x_1}}{\sigma_{x_1}}, \frac{LLD_{x_2} - \mu_{x_2}}{\sigma_{x_2}}, \sigma_{x_1x_2} \right)$$

where $BVN$, $\phi(.)$, and $\Phi(.)$ are the bivariate normal distribution, the standard univariate normal density and cumulative distribution functions, respectively, $\mu_{x_i}, \sigma_{x_i}$ is the mean of $X_i$ and the standard deviation of $X_i$, respectively; $\mu_{x_i|x_j}$ is the conditional mean of $X_i$ given $X_j$; $\sigma_{x_i|x_j}$ is the conditional standard deviation of $X_i$.
given \( X_j \) for \( i, j = 1, 2 \) and \( i \neq j \).

Let \( (x_{1i}, x_{2i}), i = 1, ..., N \) be a random sample from random variables \( (X_1, X_2) \). The maximum likelihood estimates can be obtained by maximizing the observed data likelihood, based on \( N \) pairs of \( (X_1, X_2) \). Then the likelihood function can be written as

\[
L = \prod_{i=1}^{N} \left[ \text{BVN} \left( \frac{x_{1i} - \mu_{x_1}}{\sigma_{x_1}}, \frac{x_{2i} - \mu_{x_2}}{\sigma_{x_2}}, \sigma_{x_1,x_2} \right) \right]^{d_{1i}} \left[ \Phi \left( \frac{LLD_{x_1} - \mu_{x_1}}{\sigma_{x_1}} \right) \phi \left( \frac{x_{1i} - \mu_{x_1}}{\sigma_{x_1}} \right) \right]^{d_{2i}} \times \left[ \phi \left( \frac{x_{1i} - \mu_{x_1}}{\sigma_{x_1}} \right) \Phi \left( \frac{LLD_{x_2} - \mu_{x_2}}{\sigma_{x_2}} \right) \right]^{d_{3i}} \times \left[ \text{BVN} \left( \frac{LLD_{x_1} - \mu_{x_1}}{\sigma_{x_1}}, \frac{LLD_{x_2} - \mu_{x_2}}{\sigma_{x_2}}, \sigma_{x_1,x_2} \right) \right]^{d_{4i}}
\]

where \( d_{1i}, d_{2i}, d_{3i}, d_{4i} \) are indicator variables for the following four conditions, respectively: (1) \( x_{1i} \geq LL_{x_1} \) and \( x_{2i} \geq LL_{x_2} \), (2) \( x_{1i} < LL_{x_1} \) and \( x_{2i} \geq LL_{x_2} \), (3) \( x_{1i} \geq LL_{x_1} \) and \( x_{2i} < LL_{x_2} \), (4) \( x_{1i} < LL_{x_1} \) and \( x_{2i} < LL_{x_2} \).

**Method 2: Maximum likelihood based on the bivariate Weibull-Gamma distribution**

The maximum likelihood estimates can be obtained by maximizing the observed data likelihood, based on \( N \) pairs of \( (X_1, X_2) \). All estimated CCCs \( (\hat{\rho}_c) \) were obtained by maximizing the likelihood function with respect to each of the following four scenarios:

**Case 1:** \( x_1 \geq LL_{x_1} \) and \( x_2 \geq LL_{x_2} \)

We use the bivariate Weibull distribution under the univariate cumulative Weibull distribution functions and the univariate Weibull density function.

\[
f_{12}(x_1, x_2) = f_1(x_1)f_2(x_2) \left\{ 1 + \alpha_{12} \left( g_1(H_1(x_1)) - g_1(H_1(x_1)) \right) \left( g_2(H_2(x_2)) - g_2(H_2(x_2)) \right) \right\}
\]

\[
= (\beta_1 \lambda_1 x_1^{\beta_1 - 1})(\beta_2 \lambda_2 x_2^{\beta_2 - 1}) \exp(-\lambda_1 x_1^{\beta_1} - \lambda_2 x_2^{\beta_2})
\]

\[
\times \left[ 1 + \alpha_{12} \left( 1 - \exp(-\lambda_1 x_1^{\beta_1}) - 2\lambda_1 x_1^{\beta_1} \exp(-\lambda_1 x_1^{\beta_1}) \right) \right]
\]

\[
\times \left( 1 - \exp(-\lambda_2 x_2^{\beta_2}) - 2\lambda_2 x_2^{\beta_2} \exp(-\lambda_2 x_2^{\beta_2}) \right)
\]
The maximum likelihood estimates can be obtained by maximizing the observed data likelihood, based on \( N \) pairs of \( (x_1, x_2) \).

**Case 2:** \( x_1 < \text{LLD}_{x_1} \) and \( x_2 \geq \text{LLD}_{x_2} \)

\[
\int_{0}^{\text{LLD}_{x_1}} \int_{0}^{\text{LLD}_{x_2}} f_{12}(x_1, x_2) dx_1 dx_2 = (\beta_2 \lambda_2 x_2^{\beta_2 - 1}) \exp(-\lambda_2 x_2^{\beta_2}) \left[ 1 - \exp(-\lambda_1 \text{LLD}_{x_1}^{\beta_1}) \right] \\
+ \left( \alpha_{12} \exp(-\lambda_1 \text{LLD}_{x_1}^{\beta_1}) \left( 1 - \exp(-\lambda_2 x_2^{\beta_2}) - 2\lambda_2 x_2^{\beta_2} \exp(-\lambda_2 x_2^{\beta_2}) \right) \right) \\
\times \left( -1 + \exp(-\lambda_1 \text{LLD}_{x_1}^{\beta_1}) + \lambda_1 \text{LLD}_{x_1}^{\beta_1} \exp(-\lambda_1 \text{LLD}_{x_1}^{\beta_1}) \right) \\
\times \left( -1 + \exp(-\lambda_2 x_2^{\beta_2}) + \lambda_2 \text{LLD}_{x_2}^{\beta_2} \exp(-\lambda_2 \text{LLD}_{x_2}^{\beta_2}) \right)
\]

**Case 3:** \( x_1 \geq \text{LLD}_{x_1} \) and \( x_2 < \text{LLD}_{x_2} \)

\[
\int_{0}^{\text{LLD}_{x_1}} \int_{0}^{\text{LLD}_{x_2}} f_{12}(x_1, x_2) dx_1 dx_2 = (\beta_1 \lambda_1 x_1^{\beta_1 - 1}) \exp(-\lambda_1 x_1^{\beta_1}) \left[ 1 - \exp(-\lambda_2 \text{LLD}_{x_2}^{\beta_2}) \right] \\
+ \left( \alpha_{12} \exp(-\lambda_2 \text{LLD}_{x_2}^{\beta_2}) \left( 1 - \exp(-\lambda_1 x_1^{\beta_1}) - 2\lambda_1 x_1^{\beta_1} \exp(-\lambda_1 x_1^{\beta_1}) \right) \right) \\
\times \left( -1 + \exp(-\lambda_2 x_2^{\beta_2}) + \lambda_2 \text{LLD}_{x_2}^{\beta_2} \exp(-\lambda_2 \text{LLD}_{x_2}^{\beta_2}) \right)
\]

**Case 4:** \( x_1 < \text{LLD}_{x_1} \) and \( x_2 < \text{LLD}_{x_2} \)

\[
\int_{0}^{\text{LLD}_{x_1}} \int_{0}^{\text{LLD}_{x_2}} f_{12}(x_1, x_2) dx_1 dx_2 = \left[ 1 - \exp(-\lambda_1 \text{LLD}_{x_1}^{\beta_1}) \right] \left[ 1 - \exp(-\lambda_2 \text{LLD}_{x_2}^{\beta_2}) \right] \\
+ \left[ \alpha_{12} \exp(-\lambda_1 \text{LLD}_{x_1}^{\beta_1}) \right] \\
\times \left( -1 + \exp(-\lambda_1 \text{LLD}_{x_1}^{\beta_1}) + \lambda_1 \text{LLD}_{x_1}^{\beta_1} \exp(-\lambda_1 \text{LLD}_{x_1}^{\beta_1}) \right) \\
\times \left( -1 + \exp(-\lambda_2 x_2^{\beta_2}) + \lambda_2 \text{LLD}_{x_2}^{\beta_2} \exp(-\lambda_2 \text{LLD}_{x_2}^{\beta_2}) \right)
\]

Let \( (x_{i1}, x_{i2}), i = 1, ..., N \) be a random sample from random variables \( (X_1, X_2) \). The maximum likelihood estimates can be obtained by maximizing the observed data likelihood, based on \( N \) pairs of \( (X_1, X_2) \). Then the likelihood function can be written as

\[
L = \prod_{i=1}^{N} \left[ f_{12}(x_{i1}, x_{i2}) \right]^{d_{i1}} \left[ \int_{0}^{\text{LLD}_{x_1}} f_{12}(x_{i1}, x_{i2}) dx_1 \right]^{d_{i2}} \\
\times \left[ \int_{0}^{\text{LLD}_{x_2}} f_{12}(x_{i1}, x_{i2}) dx_2 \right]^{d_{i3}} \left[ \int_{0}^{\text{LLD}_{x_1}} \int_{0}^{\text{LLD}_{x_2}} f_{12}(x_{i1}, x_{i2}) dx_2 dx_1 \right]^{d_{i4}}
\]
where $d_{1i}, d_{2i}, d_{3i}, d_{4i}$ are indicator variables for the following four conditions, respectively: (1) $x_{1i} \geq LLD_{x_1}$ and $x_{2i} \geq LLD_{x_2}$, (2) $x_{1i} < LLD_{x_1}$ and $x_{2i} \geq LLD_{x_2}$, (3) $x_{1i} \geq LLD_{x_1}$ and $x_{2i} < LLD_{x_2}$, (4) $x_{1i} < LLD_{x_1}$ and $x_{2i} < LLD_{x_2}$.

**Method 3:** Maximum likelihood based on the bivariate FGM-Weibull distribution

The maximum likelihood estimates can be obtained by maximizing the observed data likelihood, based on $N$ pairs of $(X_1, X_2)$. All estimated CCCs ($\hat{\rho}_c$) were obtained by maximizing the likelihood function with respect to each of the following four scenarios:

**Case 1:** $x_{1i} \geq LLD_{x_1}$ and $x_{2i} \geq LLD_{x_2}$ We use the bivariate Farley-Gumbel-Morgenstern density function under the univariate cumulative Weibull distribution functions and the univariate Weibull density function.

$$f_{12}(x_1, x_2) = (\beta_1 \lambda_1 x_1^{\beta_1-1}) \exp(-\lambda_1 x_1^{\beta_1})(\beta_2 \lambda_2 x_2^{\beta_2-1}) \exp(-\lambda_2 x_2^{\beta_2}) \times \left[1 + \alpha_{12} \left(1 - 2 \exp(-\lambda_1 x_1^{\beta_1})\right) \left(1 - 2 \exp(-\lambda_2 x_2^{\beta_2})\right)\right]$$

**Case 2:** $x_{1i} < LLD_{x_1}$ and $x_{2i} \geq LLD_{x_2}$

$$\int_0^{LLD_{x_1}} f_{12}(x_1, x_2)dx_1 = (\beta_2 \lambda_2 x_2^{\beta_2-1}) \exp(-\lambda_2 x_2^{\beta_2}) \left(1 - \exp(-\lambda_1 LLD_{x_1}^{\beta_1})\right) \times \left[1 + \alpha_{12} \left(2 \exp(-\lambda_2 x_2^{\beta_2}) - 1\right) \exp(-\lambda_1 LLD_{x_1}^{\beta_1})\right]$$

**Case 3:** $x_{1i} \geq LLD_{x_1}$ and $x_{2i} < LLD_{x_2}$

$$\int_0^{LLD_{x_2}} f_{12}(x_1, x_2)dx_2 = (\beta_1 \lambda_1 x_1^{\beta_1-1}) \exp(-\lambda_1 x_1^{\beta_1}) \left(1 - \exp(-\lambda_2 LLD_{x_2}^{\beta_2})\right) \times \left[1 + \alpha_{12} \left(2 \exp(-\lambda_1 x_1^{\beta_1}) - 1\right) \exp(-\lambda_2 LLD_{x_2}^{\beta_2})\right]$$
Case 4: \( x_1 < LLD_{x_1} \) and \( x_2 < LLD_{x_2} \)

\[
\int_0^{LLD_{x_1}} \int_0^{LLD_{x_2}} f_{12}(x_1, x_2)dx_2dx_1 = 
\left( 1 - \exp(-\lambda_1 LLD_{x_1}^{\beta_1}) \right) \left( 1 - \exp(-\lambda_2 LLD_{x_2}^{\beta_2}) \right) 
\times \left[ 1 + \alpha_{12} \exp(-\lambda_1 LLD_{x_1}^{\beta_1}) \exp(-\lambda_2 LLD_{x_2}^{\beta_2}) \right]
\]

Let \((x_{1i}, x_{2i}), i = 1, \ldots, N\) be a random sample from random variables \((X_1, X_2)\). The maximum likelihood estimates can be obtained by maximizing the observed data likelihood, based on \(N\) pairs of \((X_1, X_2)\). Then the likelihood function can be written as

\[
L = \prod_{i=1}^{N} \left[ f_{12}(x_{1i}, x_{2i}) \right]^{d_{1i}} \left[ \int_0^{LLD_{x_1}} f_{12}(x_{1i}, x_{2i})dx_1 \right]^{d_{2i}} 
\times \left[ \int_0^{LLD_{x_2}} f_{12}(x_{1i}, x_{2i})dx_2 \right]^{d_{3i}} \left[ \int_0^{LLD_{x_2}} \int_0^{LLD_{x_2}} f_{12}(x_{1i}, x_{2i})dx_2dx_1 \right]^{d_{4i}}
\]

where \(d_{1i}, d_{2i}, d_{3i}, d_{4i}\) are indicator variables for the following four conditions, respectively: (1) \( x_{1i} \geq LLD_{x_1} \) and \( x_{2i} \geq LLD_{x_2} \), (2) \( x_{1i} < LLD_{x_1} \) and \( x_{2i} \geq LLD_{x_2} \), (3) \( x_{1i} \geq LLD_{x_1} \) and \( x_{2i} < LLD_{x_2} \), (4) \( x_{1i} < LLD_{x_1} \) and \( x_{2i} < LLD_{x_2} \).

Remark: The bivariate FGM distribution cannot accommodate a high level of correlation. Theoretically, the alpha parameter lies within \((-1, +1)\), and correspondingly, the correlation lies within \((-0.25, +0.25)\) for many distributions used in the construction of the bivariate FGM distribution. Nevertheless, the estimated value of alpha for a specific data set could lie outside of \((-1, +1)\) and still yield a viable bivariate distribution function. Therefore, in this simulation study, we accepted estimated values of alpha within \((-4, +4)\) so that the correlation lies within \((-1, +1)\).

Method 4: Maximum likelihood based on the bivariate Piecewise Uniform-Weibull distribution

Let \(X\) be a nonnegative-valued random variable such that it is a mixture of a uniform random variable and a Weibull random variable, defined as

\[
X = I(0 \leq X \leq LLD) \cdot U + I(X > LLD) \cdot W
\]
where $LLD$ represents the lower limit of detection (a known positive value) and $I(A)$ represents the indicator function that yields 1 or 0 depending on whether the event $A$ holds. Then the density function for $X$ is

$$f(x) = \theta \cdot I(0 \leq x \leq LLD) \left[ \frac{1}{LLD} \right] + (1 - \theta) \cdot I(x > LLD) \left[ \lambda \beta (x - LLD)^{\beta - 1} \exp(-\lambda (x - LLD)^\beta) \right]$$

where $\theta = \Pr(0 \leq X \leq LLD)$.

The maximum likelihood estimates can be obtained by maximizing the observed data likelihood, based on $N$ pairs of $(X_1, X_2)$. All estimated CCCs ($\hat{\rho}_c$) were obtained by maximizing the likelihood function with respect to each of the following four scenarios:

**Case 1:** $x_1 \geq LLD_{x_1}$ and $x_2 \geq LLD_{x_2}$

$$X_1 \sim Weibull(\lambda, \beta) \quad \text{and} \quad X_2 \sim Weibull(\lambda, \beta)$$

**Case 2:** $x_1 < LLD_{x_1}$ and $x_2 \geq LLD_{x_2}$

$$X_1 \sim U[0, LLD_{x_1}] \quad \text{and} \quad X_2 \sim Weibull(\lambda, \beta)$$

**Case 3:** $x_1 \geq LLD_{x_1}$ and $x_2 < LLD_{x_2}$

$$X_1 \sim Weibull(\lambda, \beta) \quad \text{and} \quad X_2 \sim U[0, LLD_{x_2}]$$

**Case 4:** $x_1 < LLD_{x_1}$ and $x_2 < LLD_{x_2}$

$$X_1 \sim U[0, LLD_{x_1}] \quad \text{and} \quad X_2 \sim U[0, LLD_{x_2}]$$

In Tables 5.1-5.9, we report the results of a simulation study to assess the means and the standard deviations for estimating the CCC based on the ML approach under four different distributions that were aforementioned. In addition to means and standard deviations, we also report the relative bias, the mean of the standard error, and the percentage of 95\% confidence intervals (CI) that include the true value of the CCC for the 1,000 simulated data sets.
Note that the 95% confidence intervals (CI) of CCC were calculated by using the estimate statement in proc nlmixed in SAS 9.3 statistical software. Because the CCC is a function of the means, variances, and covariance under the lognormality assumption, we can derive the 95% confidence interval (CI) of the CCC under the lognormality assumption by using the delta method (See [27]).

As we expected since the data set is generated from the bivariate lognormal distribution, the ML method based on the bivariate lognormal distribution performs best among four approaches. It provides an excellent estimate of the true value of the CCC even when the censoring percentages increased, but tends to slightly underestimate the true value. The estimates from the other three distribution models are biased, although the ML method based on the bivariate FGM-Weibull distribution is clearly preferable to the ML method based on the bivariate Weibull-Gamma and the piecewise uniform-Weibull distribution for all ranges of sample sizes. Moreover, the ML method based on the bivariate FGM-Weibull distribution worked much better for most situations when the the correlations between $X$ and $Y$ are high.

From Table 5.1 (sample size of 100 paired data points, 25% left-censoring for X, 25% left-censoring for Y, and a true CCC of 0.714), the relative bias is -0.88% for the ML method based on the bivariate lognormal distribution, 13.85% for the ML method based on the bivariate Weibull-Gamma distribution, 1.04% for the ML method based on the bivariate FGM-Weibull distribution, and -73.61% for the ML method based on the bivariate Piecewise Uniform-Weibull distribution. Moreover, the ML approach based on the bivariate lognormal distribution yields the highest percentage of the 95% CI that include the true value of CCC among four distributions for all cases and for all ranges of correlation.

To see the impact of the percent of censoring, in Table 5.2 and 5.3, we increase the censoring rate to 40% and 60%, respectively. The relative biases are increased in all approaches. However, the ML approach based on the bivariate lognormal distribution still yields the smallest relative bias. From Table 5.2 (sample size of 100 paired data points, 40% left-censoring for X, 25% left-censoring for Y, and a true CCC of 0.714), the relative bias is -1.04% for the ML method based on the bivariate lognormal distribution, 28.99% for the ML method based on the bivariate Weibull-Gamma distribution, -10.66% for the ML method based on the bivariate FGM-Weibull distribution, and -73.52% for the ML method based on the bivariate Piecewise Uniform-Weibull distribution. The ML approach based on the bivariate
lognormal distribution also has the highest percentage of the 95% CI that include the true value of CCC.

Due to the large sample size of both assays (sample size = 100) in Tables 5.1-5.3, the ML method based on the bivariate lognormal distribution displays an excellent result for estimating the CCC with respect to the relative bias and the percentage of confidence intervals that include the true value of the CCC. However, if the sample size were smaller, then the ML method might produce less convincing results. To illustrate this point, we re-conduct the simulation studies with sample sizes = 50 (Tables 5.4-5.6) and sample sizes = 25 (Tables 5.7-5.9). In all cases, the ML method based on the bivariate lognormal distribution still performs best among the four approaches according to the means and the standard deviations for estimating the CCC, the relative bias, mean of the standard error, and the percentage of 95% CI that include the true value of CCC.

In general, the bivariate Weibull cannot accommodate larger correlations, which contributes to the large biases that are observed in these simulation studies.
Table 5.1. Simulation Results Based on 1000 Data Sets with Sample Size of 100 – Percent Censoring (25%, 25%)

<table>
<thead>
<tr>
<th>True $\rho$</th>
<th>True $\rho_c$</th>
<th>Distribution</th>
<th>Mean $\hat{\rho}_c$</th>
<th>Relative bias (%)</th>
<th>Empirical SD</th>
<th>Mean SE</th>
<th>The percentage of 95% confidence intervals that include true value of CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.238</td>
<td>1. Lognormal</td>
<td>0.2346</td>
<td>-1.43</td>
<td>0.0944</td>
<td>0.0939</td>
<td>94.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.0968</td>
<td>-59.33</td>
<td>0.0830</td>
<td>0.0699</td>
<td>48.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.1643</td>
<td>-30.92</td>
<td>0.0779</td>
<td>0.0656</td>
<td>75.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1427</td>
<td>-40.04</td>
<td>0.0526</td>
<td>0.0501</td>
<td>55.8</td>
</tr>
<tr>
<td>0.50</td>
<td>0.476</td>
<td>1. Lognormal</td>
<td>0.4701</td>
<td>-1.24</td>
<td>0.0807</td>
<td>0.0780</td>
<td>93.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.2404*</td>
<td>-49.50</td>
<td>0.1400</td>
<td>0.0725</td>
<td>16.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.3446</td>
<td>-27.61</td>
<td>0.1387</td>
<td>0.0630</td>
<td>29.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1861</td>
<td>-60.90</td>
<td>0.0267</td>
<td>0.0255</td>
<td>0</td>
</tr>
<tr>
<td>0.75</td>
<td>0.714</td>
<td>1. Lognormal</td>
<td>0.7077</td>
<td>-0.88</td>
<td>0.0525</td>
<td>0.0503</td>
<td>94.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.8129**</td>
<td>13.85</td>
<td>0.6151</td>
<td>0.1712</td>
<td>37.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.7214</td>
<td>1.04</td>
<td>0.1929</td>
<td>0.0848</td>
<td>61.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1884</td>
<td>-73.61</td>
<td>0.0236</td>
<td>0.0250</td>
<td>0</td>
</tr>
</tbody>
</table>

*The estimates for the bivariate Weibull-Gamma model were incalculable 1 out of 1000 data sets.

**The estimates for the bivariate Weibull-Gamma model were incalculable 38 out of 1000 data sets.
Table 5.2. Simulation Results Based on 1000 Data Sets with Sample Size of 100 – Percent Censoring (40%, 25%)

<table>
<thead>
<tr>
<th>True ρ</th>
<th>True ρc</th>
<th>Distribution</th>
<th>Mean ˆρc</th>
<th>Relative bias (%)</th>
<th>Empirical SD</th>
<th>Mean SE</th>
<th>The percentage of 95% confidence intervals that include true value of CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.238</td>
<td>1. Lognormal</td>
<td>0.2340</td>
<td>-1.68</td>
<td>0.0984</td>
<td>0.0971</td>
<td>93.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.1136</td>
<td>-52.27</td>
<td>0.1958</td>
<td>0.0837</td>
<td>56.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.1582</td>
<td>-33.53</td>
<td>0.0776</td>
<td>0.0667</td>
<td>72.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1451</td>
<td>-39.03</td>
<td>0.0523</td>
<td>0.0504</td>
<td>56.1</td>
</tr>
<tr>
<td>0.50</td>
<td>0.476</td>
<td>1. Lognormal</td>
<td>0.4692</td>
<td>-1.43</td>
<td>0.0847</td>
<td>0.0811</td>
<td>93.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.2848*</td>
<td>-40.17</td>
<td>0.3454</td>
<td>0.0899</td>
<td>25.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.3305</td>
<td>-30.57</td>
<td>0.1471</td>
<td>0.0644</td>
<td>28.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1870</td>
<td>-60.71</td>
<td>0.0261</td>
<td>0.0252</td>
<td>0</td>
</tr>
<tr>
<td>0.75</td>
<td>0.714</td>
<td>1. Lognormal</td>
<td>0.7066</td>
<td>-1.04</td>
<td>0.0554</td>
<td>0.0528</td>
<td>94.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.9210**</td>
<td>28.99</td>
<td>0.6603</td>
<td>0.1857</td>
<td>43.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.6379</td>
<td>-10.66</td>
<td>0.1986</td>
<td>0.0763</td>
<td>46.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1891</td>
<td>-73.52</td>
<td>0.0234</td>
<td>0.0237</td>
<td>0</td>
</tr>
</tbody>
</table>

*The estimates for the bivariate Weibull-Gamma model were incalculable 1 out of 1000 data sets.

**The estimates for the bivariate Weibull-Gamma model were incalculable 53 out of 1000 data sets.
Table 5.3. Simulation Results Based on 1000 Data Sets with Sample Size of 100 – Percent Censoring (60%, 25%)

<table>
<thead>
<tr>
<th>True $\rho$</th>
<th>True $\rho_c$</th>
<th>Distribution</th>
<th>Mean $\hat{\rho}_c$</th>
<th>Relative bias (%)</th>
<th>Empirical SD</th>
<th>Mean SE</th>
<th>The percentage of 95% confidence intervals that include true value of CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.238</td>
<td>1. Lognormal</td>
<td>0.2310</td>
<td>-2.94</td>
<td>0.1066</td>
<td>0.1049</td>
<td>93.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.1098</td>
<td>-53.87</td>
<td>0.0782</td>
<td>0.0711</td>
<td>59.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.1550</td>
<td>-34.87</td>
<td>0.0789</td>
<td>0.0724</td>
<td>76.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1440</td>
<td>-39.50</td>
<td>0.0517</td>
<td>0.0505</td>
<td>54.6</td>
</tr>
<tr>
<td>0.50</td>
<td>0.476</td>
<td>1. Lognormal</td>
<td>0.4656</td>
<td>-2.18</td>
<td>0.0917</td>
<td>0.0885</td>
<td>94.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.2820</td>
<td>-40.76</td>
<td>0.1308</td>
<td>0.0838</td>
<td>30.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.3265</td>
<td>-31.41</td>
<td>0.0999</td>
<td>0.0705</td>
<td>42.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1863</td>
<td>-60.86</td>
<td>0.0259</td>
<td>0.0249</td>
<td>0</td>
</tr>
<tr>
<td>0.75</td>
<td>0.714</td>
<td>1. Lognormal</td>
<td>0.7031</td>
<td>-1.53</td>
<td>0.0609</td>
<td>0.0587</td>
<td>95.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.7021*</td>
<td>-1.67</td>
<td>0.5813</td>
<td>0.153</td>
<td>36.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.5941</td>
<td>-16.79</td>
<td>0.1408</td>
<td>0.0757</td>
<td>49.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1887</td>
<td>-73.57</td>
<td>0.0233</td>
<td>0.0243</td>
<td>0</td>
</tr>
</tbody>
</table>

*The estimates for the bivariate Weibull-Gamma model were incalculable 8 out of 1000 data sets.
Table 5.4. Simulation Results Based on 1000 Data Sets with Sample Size of 50 – Percent Censoring (25%, 25%)

<table>
<thead>
<tr>
<th>True $\rho$</th>
<th>True $\rho_c$</th>
<th>Distribution</th>
<th>Mean $\hat{\rho}_c$</th>
<th>Relative bias (%)</th>
<th>Empirical SD</th>
<th>Empirical Mean $\hat{\rho}_c$</th>
<th>The percentage of 95% confidence intervals that include true value of CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.238</td>
<td>1. Lognormal</td>
<td>0.2310</td>
<td>-2.94</td>
<td>0.1351</td>
<td>0.1304</td>
<td>92.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.1039</td>
<td>-56.34</td>
<td>0.2526</td>
<td>0.1118</td>
<td>61.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.1707</td>
<td>-28.28</td>
<td>0.1348</td>
<td>0.0957</td>
<td>79.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1344</td>
<td>-43.53</td>
<td>0.0709</td>
<td>0.0676</td>
<td>71.0</td>
</tr>
<tr>
<td>0.50</td>
<td>0.476</td>
<td>1. Lognormal</td>
<td>0.4636</td>
<td>-2.61</td>
<td>0.1151</td>
<td>0.1095</td>
<td>93.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.3571*</td>
<td>-24.98</td>
<td>0.4534</td>
<td>0.1605</td>
<td>36.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.4073</td>
<td>-14.43</td>
<td>0.2265</td>
<td>0.1042</td>
<td>49.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1849</td>
<td>-61.16</td>
<td>0.0383</td>
<td>0.0373</td>
<td>0</td>
</tr>
<tr>
<td>0.75</td>
<td>0.714</td>
<td>1. Lognormal</td>
<td>0.7025</td>
<td>-1.61</td>
<td>0.0759</td>
<td>0.0716</td>
<td>93.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>1.1674 **</td>
<td>63.50</td>
<td>0.9234</td>
<td>0.3297</td>
<td>36.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.7528</td>
<td>5.43</td>
<td>0.2001</td>
<td>0.1182</td>
<td>70.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1898</td>
<td>-73.42</td>
<td>0.0338</td>
<td>0.0340</td>
<td>0</td>
</tr>
</tbody>
</table>

*The estimates for the bivariate Weibull-Gamma model were incalculable 3 out of 1000 data sets.

**The estimates for the bivariate Weibull-Gamma model were incalculable 153 out of 1000 data sets.
Table 5.5. Simulation Results Based on 1000 Data Sets with Sample Size of 50 – Percent Censoring (40%, 25%)

<table>
<thead>
<tr>
<th>True ρ</th>
<th>True ρc</th>
<th>Distribution</th>
<th>Mean $\hat{\rho}_c$</th>
<th>Relative bias (%)</th>
<th>Empirical SD</th>
<th>Empirical SE</th>
<th>The percentage of 95% confidence intervals that include true value of CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.238</td>
<td>1. Lognormal</td>
<td>0.2296</td>
<td>-3.53</td>
<td>0.1395</td>
<td>0.1345</td>
<td>92.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull - Gamma</td>
<td>0.1179</td>
<td>-50.46</td>
<td>0.2844</td>
<td>0.1214</td>
<td>68.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM - Weibull</td>
<td>0.1628</td>
<td>-31.60</td>
<td>0.1228</td>
<td>0.0955</td>
<td>80.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1363</td>
<td>-42.73</td>
<td>0.0710</td>
<td>0.0677</td>
<td>71.5</td>
</tr>
<tr>
<td>0.50</td>
<td>0.476</td>
<td>1. Lognormal</td>
<td>0.4617</td>
<td>-3.00</td>
<td>0.1206</td>
<td>0.1135</td>
<td>93.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull - Gamma</td>
<td>0.3768*</td>
<td>-20.84</td>
<td>0.4513</td>
<td>0.1586</td>
<td>41.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM - Weibull</td>
<td>0.3700</td>
<td>-22.27</td>
<td>0.2010</td>
<td>0.0981</td>
<td>48.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1858</td>
<td>-60.97</td>
<td>0.0377</td>
<td>0.0369</td>
<td>0.1</td>
</tr>
<tr>
<td>0.75</td>
<td>0.714</td>
<td>1. Lognormal</td>
<td>0.7003</td>
<td>-1.92</td>
<td>0.0805</td>
<td>0.0750</td>
<td>93.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull - Gamma</td>
<td>1.2133**</td>
<td>69.93</td>
<td>0.8829</td>
<td>0.3627</td>
<td>39.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM - Weibull</td>
<td>0.6926</td>
<td>-3.00</td>
<td>0.2153</td>
<td>0.1135</td>
<td>60.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1906</td>
<td>-73.31</td>
<td>0.0336</td>
<td>0.0336</td>
<td>0</td>
</tr>
</tbody>
</table>

*The estimates for the bivariate Weibull-Gamma model were incalculable 22 out of 1000 data sets.  
**The estimates for the bivariate Weibull-Gamma model were incalculable 163 out of 1000 data sets.
Table 5.6. Simulation Results Based on 1000 Data Sets with Sample Size of 50 – Percent Censoring (60%, 25%)

<table>
<thead>
<tr>
<th>True ρ</th>
<th>True ρc</th>
<th>Distribution</th>
<th>Mean $\hat{\rho}_c$</th>
<th>Relative bias (%)</th>
<th>Empirical Mean SD</th>
<th>Empirical Mean SE</th>
<th>The percentage of 95% confidence intervals that include true value of CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.238</td>
<td>1. Lognormal</td>
<td>0.2257</td>
<td>-5.17</td>
<td>0.1514</td>
<td>0.1448</td>
<td>93.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull - Gamma</td>
<td>0.1174*</td>
<td>-50.67</td>
<td>0.2336</td>
<td>0.1264</td>
<td>73.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM - Weibull</td>
<td>0.1565</td>
<td>-34.24</td>
<td>0.1194</td>
<td>0.1029</td>
<td>82.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1350</td>
<td>-43.28</td>
<td>0.0706</td>
<td>0.0675</td>
<td>69.4</td>
</tr>
<tr>
<td>0.50</td>
<td>0.476</td>
<td>1. Lognormal</td>
<td>0.4571</td>
<td>-3.97</td>
<td>0.1311</td>
<td>0.1237</td>
<td>92.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull - Gamma</td>
<td>0.3214**</td>
<td>-32.48</td>
<td>0.2991</td>
<td>0.1451</td>
<td>48.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM - Weibull</td>
<td>0.3485</td>
<td>-26.79</td>
<td>0.1617</td>
<td>0.1060</td>
<td>60.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1850</td>
<td>-61.13</td>
<td>0.0378</td>
<td>0.0368</td>
<td>0.3</td>
</tr>
<tr>
<td>0.75</td>
<td>0.714</td>
<td>1. Lognormal</td>
<td>0.6961</td>
<td>-2.51</td>
<td>0.0886</td>
<td>0.0840</td>
<td>93.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull - Gamma</td>
<td>0.7809***</td>
<td>9.37</td>
<td>0.7597</td>
<td>0.2889</td>
<td>43.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM - Weibull</td>
<td>0.6372</td>
<td>-10.76</td>
<td>0.1789</td>
<td>0.1130</td>
<td>69.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1901</td>
<td>-73.38</td>
<td>0.0336</td>
<td>0.0333</td>
<td>0</td>
</tr>
</tbody>
</table>

*The estimates for the bivariate Weibull-Gamma model were incalculable 1 out of 1000 data sets.
**The estimates for the bivariate Weibull-Gamma model were incalculable 5 out of 1000 data sets.
***The estimates for the bivariate Weibull-Gamma model were incalculable 23 out of 1000 data sets.
Table 5.7. Simulation Results Based on 1000 Data Sets with Sample Size of 25 – Percent Censoring (25%, 25%)

<table>
<thead>
<tr>
<th>True $\rho$</th>
<th>True $\rho_c$</th>
<th>Distribution</th>
<th>Mean $\hat{\rho}_c$</th>
<th>Relative bias (%)</th>
<th>Empirical SD</th>
<th>Empirical Mean SE</th>
<th>The percentage of 95% confidence intervals that include true value of CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.238</td>
<td>1. Lognormal</td>
<td>0.2225</td>
<td>-6.51</td>
<td>0.1905</td>
<td>0.1783</td>
<td>93.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull - Gamma</td>
<td>0.0795*</td>
<td>-66.6</td>
<td>0.4437</td>
<td>0.2144</td>
<td>75.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM - Weibull</td>
<td>0.1961</td>
<td>-17.61</td>
<td>0.2317</td>
<td>0.1402</td>
<td>80.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1168</td>
<td>-50.92</td>
<td>0.0961</td>
<td>0.0902</td>
<td>77.6</td>
</tr>
<tr>
<td>0.50</td>
<td>0.476</td>
<td>1. Lognormal</td>
<td>0.4496</td>
<td>-5.55</td>
<td>0.1650</td>
<td>0.1526</td>
<td>93.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull - Gamma</td>
<td>0.4618**</td>
<td>-2.98</td>
<td>0.7202</td>
<td>0.2768</td>
<td>50.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM - Weibull</td>
<td>0.4578</td>
<td>-3.82</td>
<td>0.2965</td>
<td>0.1535</td>
<td>60.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1800</td>
<td>-62.18</td>
<td>0.0539</td>
<td>0.0560</td>
<td>2.8</td>
</tr>
<tr>
<td>0.75</td>
<td>0.714</td>
<td>1. Lognormal</td>
<td>0.6897</td>
<td>-3.40</td>
<td>0.1136</td>
<td>0.1027</td>
<td>92.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull - Gamma</td>
<td>1.2231***</td>
<td>71.30</td>
<td>1.0526</td>
<td>0.4972</td>
<td>38.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM - Weibull</td>
<td>0.7605</td>
<td>6.51</td>
<td>0.2356</td>
<td>0.1593</td>
<td>71.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1929</td>
<td>-72.98</td>
<td>0.0436</td>
<td>0.0455</td>
<td>0</td>
</tr>
</tbody>
</table>

*The estimates for the bivariate Weibull-Gamma model were incalculable 11 out of 1000 data sets.

**The estimates for the bivariate Weibull-Gamma model were incalculable 64 out of 1000 data sets.

***The estimates for the bivariate Weibull-Gamma model were incalculable 360 out of 1000 data sets.
Table 5.8. Simulation Results Based on 1000 Data Sets with Sample Size of 25 – Percent Censoring (40%, 25%)

<table>
<thead>
<tr>
<th>True ρ</th>
<th>True ρc</th>
<th>Distribution</th>
<th>Mean $\hat{\rho}_c$</th>
<th>Relative bias (%)</th>
<th>Empirical SD</th>
<th>Empirical SE</th>
<th>The percentage of 95% confidence intervals that include true value of CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.238</td>
<td>1. Lognormal</td>
<td>0.2193</td>
<td>-7.86</td>
<td>0.1948</td>
<td>0.1824</td>
<td>93.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull - Gamma</td>
<td>0.0809*</td>
<td>-66.01</td>
<td>0.4457</td>
<td>0.2182</td>
<td>75.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM - Weibull</td>
<td>0.1800</td>
<td>-24.37</td>
<td>0.2117</td>
<td>0.1381</td>
<td>81.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1172</td>
<td>-50.76</td>
<td>0.0966</td>
<td>0.0912</td>
<td>78.9</td>
</tr>
<tr>
<td>0.50</td>
<td>0.476</td>
<td>1. Lognormal</td>
<td>0.4449</td>
<td>-6.53</td>
<td>0.1693</td>
<td>0.1570</td>
<td>93.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull - Gamma</td>
<td>0.5089**</td>
<td>6.91</td>
<td>0.7687</td>
<td>0.3046</td>
<td>54.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM - Weibull</td>
<td>0.4186</td>
<td>-12.06</td>
<td>0.2789</td>
<td>0.1507</td>
<td>63.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1813</td>
<td>-61.91</td>
<td>0.0527</td>
<td>0.0561</td>
<td>3.7</td>
</tr>
<tr>
<td>0.75</td>
<td>0.714</td>
<td>1. Lognormal</td>
<td>0.6847</td>
<td>-4.10</td>
<td>0.1180</td>
<td>0.1074</td>
<td>92.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull - Gamma</td>
<td>1.2794***</td>
<td>79.19</td>
<td>1.0364</td>
<td>0.5819</td>
<td>40.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM - Weibull</td>
<td>0.7213</td>
<td>1.02</td>
<td>0.2422</td>
<td>0.1586</td>
<td>70.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1936</td>
<td>-72.89</td>
<td>0.0433</td>
<td>0.0452</td>
<td>0</td>
</tr>
</tbody>
</table>

*The estimates for the bivariate Weibull-Gamma model were incalculable 15 out of 1000 data sets.
**The estimates for the bivariate Weibull-Gamma model were incalculable 59 out of 1000 data sets.
***The estimates for the bivariate Weibull-Gamma model were incalculable 347 out of 1000 data sets.
Table 5.9. Simulation Results Based on 1000 Data Sets with Sample Size of 25 – Percent Censoring (60%, 25%)
5.2 Data Set Generated by a FGM-Weibull Distribution

In this section, the paired samples were generated from the bivariate FGM-Weibull distribution using SAS/IML software for each of the following four methods:

1. Maximum likelihood method based on the bivariate lognormal distribution
2. Maximum likelihood method based on the bivariate Weibull-Gamma distribution
3. Maximum likelihood method based on the bivariate FGM-Weibull distribution
4. Maximum likelihood method based on the Piecewise Uniform-Weibull distribution

We generate the paired data represented by the variables $X_1$ and $X_2$ with a sample size of 100, 50, 25 for each of 1000 data sets using one of the following parameter settings: $\lambda_1 = \beta_1 = \lambda_2 = \beta_2 = 1$, the correlation parameter $\alpha = 0.25, 0.50, 1.00$, and the correlation coefficient $\rho = 0.0625, 0.125, 0.25$, and left-censoring rates of (25\% for $X_1$, 25\% for $X_2$) or (40\% for $X_1$, 25\% for $X_2$) or (60\% for $X_1$, 25\% for $X_2$).

Let $X_1$ and $X_2$ be continuous random variables representing the two assay readings on the same subject based on two different techniques. Define $LLD_{x_1}$ and $LLD_{x_2}$ as the left-censored variables corresponding to $X_1$ and $X_2$. Let $(x_{1i}, x_{2i}), i = 1, ..., N$ be a random sample from random variables $(X_{1i}, X_{2i})$. The maximum likelihood estimates can be obtained by maximizing the observed data likelihood, based on N pairs of $(X_1, X_2)$. As in section 5.1, all estimated CCCs ($\hat{\rho}_c$) were obtained by maximizing the likelihood function with respect to each of the following four scenarios:

- **Case 1:** $x_1 \geq LLD_{x_1}$ and $x_2 \geq LLD_{x_2}$
- **Case 2:** $x_1 < LLD_{x_1}$ and $x_2 \geq LLD_{x_2}$
- **Case 3:** $x_1 \geq LLD_{x_1}$ and $x_2 < LLD_{x_2}$
- **Case 4:** $x_1 < LLD_{x_1}$ and $x_2 < LLD_{x_2}$

In Tables 5.10-5.18, we report the results of a simulation study to assess the means and the standard deviations for estimating the CCC based on the ML approach.
under four different distributions that was aforementioned. In addition to means and standard deviations, we also report the relative bias, the mean of the standard error, and the percentage of 95% confidence intervals (CI) that include the true value of CCC for the 1,000 simulated data sets.

As we expected since the data set is generated from the bivariate FGM-Weibull distribution, the ML method based on the bivariate FGM-Weibull distribution performs best among four approaches. It provides an excellent estimate of the true value of CCC even when the censoring percentages increased, but tends to slightly underestimate the true value. The estimates from other three distributions are biased, although the ML method based on the bivariate piecewise uniform-Weibull distribution is clearly preferable to the ML method based on the bivariate Weibull-Gamma and lognormal when sample sizes equal to 100. However, for the smaller sample size (sample size = 50 and 25), the ML method based on the bivariate lognormal distribution is preferable to the ML method based on the bivariate Weibull-Gamma and piecewise uniform-Weibull distribution.

From Table 5.10 (sample size of 100 paired data points, 25% left-censoring for X, 25% left-censoring for Y, and a true CCC of 0.25), the relative bias is 15.52% for the ML method based on the bivariate lognormal distribution, -37.26% for the ML method based on the bivariate Weibull-Gamma distribution, -0.16% for the ML method based on the bivariate FGM-Weibull distribution, and -11.61% for the ML method based on the bivariate piecewise uniform-Weibull distribution. All approaches yield very high percentage of the 95% CI that include the true value of CCC for all cases and for all ranges of correlation.

To see the impact of the percent of censoring, in Table 5.11 and 5.12, we increase the censoring rate to 40% and 60%, respectively. The relative biases are increased in all approaches when the censoring rate equals to 40%, but it is smaller when the censoring rate equals to 60%. However, the ML approach based on the bivariate FGM-Weibull distribution still yields the smallest relative bias. From Table 5.11 (sample size of 100 paired data points, 40% left-censoring for X, 25% left-censoring for Y, and a true CCC of 0.25), the relative bias is 16.27% for the ML method based on the bivariate lognormal distribution, -32.49% for the ML method based on the bivariate Weibull-Gamma distribution, -16.27% for the ML method based on the bivariate FGM-Weibull distribution, and -12.39% for the ML method based on the bivariate piecewise uniform-Weibull distribution. All approaches yield very high
percentage of the 95% CI that include the true value of CCC for all cases.

Due to the large sample size of both assays (sample size = 100) in Tables 5.10-5.12, the ML method based on the bivariate FGM-Weibull distribution displays an excellent result for estimating the CCC with respect to the relative bias and the percentage of confidence intervals that include the true value of the CCC. However, if the sample size were smaller, then the ML method might produce less convincing results. To illustrate this point, we re-conduct the simulation studies with sample sizes = 50 (Tables 5.13-5.15) and sample sizes = 25 (Tables 5.16-5.18). In all cases, the ML method based on the bivariate FGM-Weibull distribution still performs best among the four approaches according to the means and the standard deviations for estimating the CCC, the relative bias, mean of the standard error, and the percentage of 95% CI that include the true value of CCC.

Note that, in case of very low agreement and small sample size (a true CCC of 0.0625 and sample size = 50 and 25), the ML method based on the bivariate lognormal distribution performs better than the ML method based on the bivariate FGM-Weibull distribution based on the relative bias.
<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>True $\rho_c$</th>
<th>Distribution</th>
<th>Mean $\rho_c$</th>
<th>Relative bias (%)</th>
<th>Empirical SD</th>
<th>Mean SE</th>
<th>The percentage of 95% confidence intervals that include true value of CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.0625</td>
<td>1. Normal</td>
<td>0.0734</td>
<td>17.40</td>
<td>0.1049</td>
<td>0.1028</td>
<td>94.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.0296</td>
<td>-52.70</td>
<td>0.1124</td>
<td>0.0920</td>
<td>91.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.0601</td>
<td>-3.89</td>
<td>0.0764</td>
<td>0.0728</td>
<td>93.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.0594</td>
<td>-4.91</td>
<td>0.0736</td>
<td>0.0706</td>
<td>93.7</td>
</tr>
<tr>
<td>0.50</td>
<td>0.125</td>
<td>1. Normal</td>
<td>0.1469</td>
<td>17.55</td>
<td>0.1010</td>
<td>0.1007</td>
<td>94.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.0724</td>
<td>-42.09</td>
<td>0.1053</td>
<td>0.0914</td>
<td>90.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.1213</td>
<td>-2.96</td>
<td>0.0746</td>
<td>0.0709</td>
<td>93.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1189</td>
<td>-4.85</td>
<td>0.0692</td>
<td>0.0668</td>
<td>91.8</td>
</tr>
<tr>
<td>1.00</td>
<td>0.25</td>
<td>1. Normal</td>
<td>0.2888</td>
<td>15.52</td>
<td>0.0905</td>
<td>0.0940</td>
<td>92.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.1568</td>
<td>-37.26</td>
<td>0.1031</td>
<td>0.0880</td>
<td>80.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.2496</td>
<td>-0.16</td>
<td>0.0737</td>
<td>0.0635</td>
<td>93.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.2210</td>
<td>-11.61</td>
<td>0.0388</td>
<td>0.0393</td>
<td>97.1</td>
</tr>
</tbody>
</table>
Table 5.11. Simulation Results Based on 1000 Data Sets with Sample Size of 100 – Percent Censoring (40%, 25%)

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$\rho_c$</th>
<th>True Distribution</th>
<th>Mean $\hat{\rho}_c$</th>
<th>Relative bias (%)</th>
<th>Empirical SD</th>
<th>Mean SE</th>
<th>The percentage of 95% confidence intervals that include true value of CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.0625</td>
<td>1. Lognormal</td>
<td>0.0736</td>
<td>17.81</td>
<td>0.1062</td>
<td>0.1052</td>
<td>94.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.0315</td>
<td>-49.55</td>
<td>0.1136</td>
<td>0.0929</td>
<td>91.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.0599</td>
<td>-4.15</td>
<td>0.0782</td>
<td>0.0748</td>
<td>93.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.0584</td>
<td>-6.50</td>
<td>0.0724</td>
<td>0.0694</td>
<td>93.8</td>
</tr>
<tr>
<td>0.50</td>
<td>0.125</td>
<td>1. Lognormal</td>
<td>0.1481</td>
<td>18.49</td>
<td>0.1022</td>
<td>0.1031</td>
<td>94.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.0771</td>
<td>-38.30</td>
<td>0.1066</td>
<td>0.0929</td>
<td>90.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.1212</td>
<td>-3.05</td>
<td>0.0768</td>
<td>0.0733</td>
<td>93.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1170</td>
<td>-6.41</td>
<td>0.0684</td>
<td>0.0662</td>
<td>92.5</td>
</tr>
<tr>
<td>1.00</td>
<td>0.25</td>
<td>1. Lognormal</td>
<td>0.2907</td>
<td>16.27</td>
<td>0.0914</td>
<td>0.0961</td>
<td>92.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.1688</td>
<td>-32.49</td>
<td>0.1046</td>
<td>0.0903</td>
<td>84.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.2470</td>
<td>-1.18</td>
<td>0.0721</td>
<td>0.0665</td>
<td>93.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.2190</td>
<td>-12.39</td>
<td>0.0397</td>
<td>0.0401</td>
<td>96.3</td>
</tr>
</tbody>
</table>
Table 5.12. Simulation Results Based on 1000 Data Sets with Sample Size of 100 – Percent Censoring (60%, 25%)

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>True $\rho_c$</th>
<th>Distribution</th>
<th>Mean $\hat{\rho}_c$</th>
<th>Relative bias (%)</th>
<th>Empirical SD</th>
<th>Mean SE</th>
<th>The percentage of 95% confidence intervals that include true value of CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.0625</td>
<td>1. Lognormal</td>
<td>0.0689</td>
<td>10.19</td>
<td>0.1088</td>
<td>0.1075</td>
<td>94.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.0352</td>
<td>-43.63</td>
<td>0.1021</td>
<td>0.0882</td>
<td>91.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.0615</td>
<td>-1.54</td>
<td>0.0857</td>
<td>0.0820</td>
<td>93.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.0566</td>
<td>-9.47</td>
<td>0.0699</td>
<td>0.0673</td>
<td>93.7</td>
</tr>
<tr>
<td>0.50</td>
<td>0.125</td>
<td>1. Lognormal</td>
<td>0.1384</td>
<td>10.70</td>
<td>0.1041</td>
<td>0.1053</td>
<td>94.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.0789</td>
<td>-36.90</td>
<td>0.0995</td>
<td>0.0869</td>
<td>86.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.1231</td>
<td>-1.49</td>
<td>0.0856</td>
<td>0.0801</td>
<td>93.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1133</td>
<td>-9.32</td>
<td>0.0665</td>
<td>0.0649</td>
<td>92.9</td>
</tr>
<tr>
<td>1.00</td>
<td>0.25</td>
<td>1. Lognormal</td>
<td>0.2707</td>
<td>8.26</td>
<td>0.0929</td>
<td>0.0990</td>
<td>94.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.1593</td>
<td>-36.30</td>
<td>0.0815</td>
<td>0.0719</td>
<td>87.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.2507</td>
<td>0.26</td>
<td>0.0840</td>
<td>0.0750</td>
<td>92.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.2159</td>
<td>-13.65</td>
<td>0.0413</td>
<td>0.0424</td>
<td>96.4</td>
</tr>
</tbody>
</table>
Table 5.13. Simulation Results Based on 1000 Data Sets with Sample Size of 50 – Percent Censoring (25%, 25%)

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$\rho_c$</th>
<th>True Distribution</th>
<th>Mean $\rho_c$</th>
<th>Relative Bias (%)</th>
<th>Empirical SD</th>
<th>Mean SE</th>
<th>The percentage of 95% confidence intervals that include true value of CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.0625</td>
<td>1. Lognormal</td>
<td>0.0674</td>
<td>7.88</td>
<td>0.1474</td>
<td>0.1417</td>
<td>93.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>-0.0039</td>
<td>-106.18</td>
<td>0.2239</td>
<td>0.1475</td>
<td>91.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.0544</td>
<td>-12.89</td>
<td>0.1072</td>
<td>0.0992</td>
<td>94.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.0536</td>
<td>-14.19</td>
<td>0.1014</td>
<td>0.0959</td>
<td>93.1</td>
</tr>
<tr>
<td>0.50</td>
<td>0.125</td>
<td>1. Lognormal</td>
<td>0.1418</td>
<td>13.42</td>
<td>0.1429</td>
<td>0.1391</td>
<td>93.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.0513</td>
<td>-58.92</td>
<td>0.2056</td>
<td>0.1437</td>
<td>90.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.1155</td>
<td>-7.56</td>
<td>0.1088</td>
<td>0.0979</td>
<td>93.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1104</td>
<td>-11.70</td>
<td>0.0945</td>
<td>0.0887</td>
<td>85.6</td>
</tr>
<tr>
<td>1.00</td>
<td>0.25</td>
<td>1. Lognormal</td>
<td>0.2842</td>
<td>13.68</td>
<td>0.1286</td>
<td>0.1301</td>
<td>92.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.1681</td>
<td>-32.76</td>
<td>0.2083</td>
<td>0.1394</td>
<td>86.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.2529</td>
<td>1.17</td>
<td>0.1358</td>
<td>0.0943</td>
<td>91.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.2059</td>
<td>-17.62</td>
<td>0.0589</td>
<td>0.0568</td>
<td>96.3</td>
</tr>
</tbody>
</table>
Table 5.14. Simulation Results Based on 1000 Data Sets with Sample Size of 50 – Percent Censoring (40%, 25%)

<table>
<thead>
<tr>
<th>α</th>
<th>True Distribution</th>
<th>0.25</th>
<th>0.50</th>
<th>1.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>ρc</td>
<td></td>
<td>Mean</td>
<td>Relative bias</td>
<td>Empirical SD</td>
</tr>
<tr>
<td>0.0625</td>
<td>1. Lognormal</td>
<td>0.0690</td>
<td>10.42</td>
<td>0.1474</td>
</tr>
<tr>
<td></td>
<td>2. Weibull-Gamma</td>
<td>-0.0052</td>
<td>-108.34</td>
<td>0.2294</td>
</tr>
<tr>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.0556</td>
<td>-11.02</td>
<td>0.1074</td>
</tr>
<tr>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.0529</td>
<td>-15.34</td>
<td>0.1001</td>
</tr>
<tr>
<td>0.125</td>
<td>1. Lognormal</td>
<td>0.1441</td>
<td>-15.25</td>
<td>0.1424</td>
</tr>
<tr>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.0533</td>
<td>-57.33</td>
<td>2096</td>
</tr>
<tr>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.1164</td>
<td>-6.89</td>
<td>0.1080</td>
</tr>
<tr>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1090</td>
<td>-12.79</td>
<td>0.0938</td>
</tr>
<tr>
<td>0.25</td>
<td>1. Lognormal</td>
<td>0.2867</td>
<td>14.67</td>
<td>0.1280</td>
</tr>
<tr>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.1746</td>
<td>-30.16</td>
<td>0.1999</td>
</tr>
<tr>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.2484</td>
<td>-0.65</td>
<td>0.1257</td>
</tr>
<tr>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.2047</td>
<td>-18.14</td>
<td>0.0596</td>
</tr>
</tbody>
</table>
Table 5.15. Simulation Results Based on 1000 Data Sets with Sample Size of 50 – Percent Censoring (60%, 25%)

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$\rho_c$</th>
<th>True Distribution</th>
<th>Mean $\hat{\rho}_c$</th>
<th>Relative bias (%)</th>
<th>Empirical SD</th>
<th>Mean SE</th>
<th>The percentage of 95% confidence intervals that include true value of CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.0625</td>
<td>1. Lognormal</td>
<td>0.0637</td>
<td>1.95</td>
<td>0.1534</td>
<td>0.1499</td>
<td>94.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>-0.0068</td>
<td>-110.94</td>
<td>0.2301</td>
<td>0.1472</td>
<td>91.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.0597</td>
<td>-4.41</td>
<td>0.1222</td>
<td>0.1132</td>
<td>93.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.0512</td>
<td>-18.04</td>
<td>0.0971</td>
<td>0.0928</td>
<td>94.1</td>
</tr>
<tr>
<td>0.50</td>
<td>0.125</td>
<td>1. Lognormal</td>
<td>0.1346</td>
<td>7.66</td>
<td>0.1479</td>
<td>0.1470</td>
<td>94.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.0534</td>
<td>-57.24</td>
<td>0.2076</td>
<td>0.1435</td>
<td>90.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.1222</td>
<td>-2.23</td>
<td>0.1264</td>
<td>0.1126</td>
<td>93.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1060</td>
<td>-15.19</td>
<td>0.0915</td>
<td>0.0874</td>
<td>88.1</td>
</tr>
<tr>
<td>1.00</td>
<td>0.25</td>
<td>1. Lognormal</td>
<td>0.2678</td>
<td>7.10</td>
<td>0.1332</td>
<td>0.1383</td>
<td>93.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.1746</td>
<td>-30.15</td>
<td>0.1945</td>
<td>0.1389</td>
<td>88.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.2559</td>
<td>2.36</td>
<td>0.1417</td>
<td>0.1101</td>
<td>92.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.2015</td>
<td>-19.41</td>
<td>0.0602</td>
<td>0.0607</td>
<td>96.3</td>
</tr>
</tbody>
</table>
Table 5.16. Simulation Results Based on 1000 Data Sets with Sample Size of 25 – Percent Censoring (25%, 25%)

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>True $\rho_c$</th>
<th>Distribution</th>
<th>Mean $\hat{\rho}_c$</th>
<th>Relative bias (%)</th>
<th>Empirical SD</th>
<th>Mean SE</th>
<th>The percentage of 95% confidence intervals that include true value of CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.0625</td>
<td>1. Lognormal</td>
<td>0.0621</td>
<td>-0.71</td>
<td>0.2026</td>
<td>0.1913</td>
<td>92.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.0072</td>
<td>-88.54</td>
<td>0.1701</td>
<td>0.1299</td>
<td>57.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.0590</td>
<td>-5.53</td>
<td>0.1871</td>
<td>0.1367</td>
<td>93.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.0474</td>
<td>-24.09</td>
<td>0.1327</td>
<td>0.1205</td>
<td>82.4</td>
</tr>
<tr>
<td>0.50</td>
<td>0.125</td>
<td>1. Lognormal</td>
<td>0.1360</td>
<td>8.79</td>
<td>0.1956</td>
<td>0.1883</td>
<td>92.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.0440</td>
<td>-64.78</td>
<td>0.1678</td>
<td>0.1246</td>
<td>60.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.1244</td>
<td>-0.45</td>
<td>0.2021</td>
<td>0.1382</td>
<td>91.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.0964</td>
<td>-22.90</td>
<td>0.1234</td>
<td>0.1130</td>
<td>79.7</td>
</tr>
<tr>
<td>1.00</td>
<td>0.25</td>
<td>1. Lognormal</td>
<td>0.2793</td>
<td>11.71</td>
<td>0.1773</td>
<td>0.1768</td>
<td>92.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.1054</td>
<td>-57.85</td>
<td>0.1525</td>
<td>0.1068</td>
<td>87.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.2840</td>
<td>13.62</td>
<td>0.2495</td>
<td>0.1443</td>
<td>86.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1839</td>
<td>-26.46</td>
<td>0.0863</td>
<td>0.0823</td>
<td>96.3</td>
</tr>
</tbody>
</table>
Table 5.17. Simulation Results Based on 1000 Data Sets with Sample Size of 25 – Percent Censoring (40%, 25%)

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>True $\rho_c$</th>
<th>Distribution</th>
<th>Mean $\hat{\rho}_c$</th>
<th>Relative bias (%)</th>
<th>Empirical SD</th>
<th>Mean SE</th>
<th>The percentage of 95% confidence intervals that include true value of CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.0625</td>
<td>1. Lognormal</td>
<td>0.0635</td>
<td>1.52</td>
<td>0.2037</td>
<td>0.1956</td>
<td>93.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.0075</td>
<td>-87.96</td>
<td>0.1699</td>
<td>0.1252</td>
<td>56.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.0592</td>
<td>-5.21</td>
<td>0.1793</td>
<td>0.1404</td>
<td>93.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.0471</td>
<td>-24.62</td>
<td>0.1318</td>
<td>0.1189</td>
<td>82.7</td>
</tr>
<tr>
<td>0.50</td>
<td>0.125</td>
<td>1. Lognormal</td>
<td>0.1375</td>
<td>9.96</td>
<td>0.1971</td>
<td>0.1924</td>
<td>93.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.0442</td>
<td>-64.67</td>
<td>0.1669</td>
<td>0.1241</td>
<td>60.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.1221</td>
<td>-2.28</td>
<td>0.1924</td>
<td>0.1408</td>
<td>92.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.0958</td>
<td>-23.35</td>
<td>0.1223</td>
<td>0.1128</td>
<td>81.3</td>
</tr>
<tr>
<td>1.00</td>
<td>0.25</td>
<td>1. Lognormal</td>
<td>0.2807</td>
<td>12.30</td>
<td>0.1791</td>
<td>0.1805</td>
<td>93.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.1093</td>
<td>-56.30</td>
<td>0.1506</td>
<td>0.1072</td>
<td>87.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.2731</td>
<td>9.25</td>
<td>0.2364</td>
<td>0.1447</td>
<td>88.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1829</td>
<td>-26.86</td>
<td>0.0860</td>
<td>0.0831</td>
<td>95.0</td>
</tr>
</tbody>
</table>
Table 5.18. Simulation Results Based on 1000 Data Sets with Sample Size of 25 – Percent Censoring (60%, 25%)

<table>
<thead>
<tr>
<th>α</th>
<th>True ρc</th>
<th>Distribution</th>
<th>Mean ρc</th>
<th>Relative bias (%)</th>
<th>Empirical SD</th>
<th>Empirical Mean SE</th>
<th>The percentage of 95% confidence intervals that include true value of CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.0625</td>
<td>1. Lognormal</td>
<td>0.0583</td>
<td>-6.79</td>
<td>0.2140</td>
<td>0.2026</td>
<td>93.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.0053</td>
<td>-91.51</td>
<td>0.1677</td>
<td>0.1239</td>
<td>57.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.0664</td>
<td>6.21</td>
<td>0.2035</td>
<td>0.1572</td>
<td>94.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.0457</td>
<td>26.90</td>
<td>0.1288</td>
<td>0.1182</td>
<td>84.7</td>
</tr>
<tr>
<td>0.50</td>
<td>0.125</td>
<td>1. Lognormal</td>
<td>0.1280</td>
<td>2.39</td>
<td>0.2077</td>
<td>0.1993</td>
<td>93.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.0428</td>
<td>-65.75</td>
<td>0.1647</td>
<td>0.1224</td>
<td>61.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.1322</td>
<td>5.74</td>
<td>0.2157</td>
<td>0.1594</td>
<td>93.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.0939</td>
<td>-24.86</td>
<td>0.1201</td>
<td>0.1128</td>
<td>83.4</td>
</tr>
<tr>
<td>1.00</td>
<td>0.25</td>
<td>1. Lognormal</td>
<td>0.2628</td>
<td>5.10</td>
<td>0.1892</td>
<td>0.1886</td>
<td>93.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Weibull-Gamma</td>
<td>0.1083</td>
<td>-56.69</td>
<td>0.1481</td>
<td>0.1070</td>
<td>87.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. FGM-Weibull</td>
<td>0.2824</td>
<td>12.97</td>
<td>0.2454</td>
<td>0.1652</td>
<td>89.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Piecewise Uniform-Weibull</td>
<td>0.1814</td>
<td>-27.45</td>
<td>0.0860</td>
<td>0.0859</td>
<td>94.6</td>
</tr>
</tbody>
</table>
Chapter 6
Examples

In this chapter, we demonstrate the use of the ML method based on a new class of bivariate Weibull distributions, presented in Chapter 4, to find the CCC for measuring overall agreement between two variables with left-censoring in the first example and right-censoring in the second example.

6.1 Urine Stability Studies for Novel Biomarkers of Acute Kidney Injury

The data for this example are taken from the Assessment, Serial Evaluation, and Subsequent Sequelae in Acute Kidney Injury Study (ASSESS-AKI), which is a National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)-sponsored multi-site research consortium. This data set was first analyzed by Parikh et. al [2]. The main objective of this trial was to determine the agreement between the measurements of the urinary biomarkers collected under a standard condition and values obtained in samples prepared under three different processes, labeled as Process A, Process B, and Process C.

- Process A: Centrifugation followed by immediate aliquoting and temporary storage at \(+4^\circ\text{C}\) for 48 hours prior to freezing at \(-80^\circ\text{C}\)
- Process B: Centrifugation followed by immediate aliquoting and temporary storage at \(+25^\circ\text{C}\) for 48 hours prior to freezing at \(-80^\circ\text{C}\)
- Process C: Samples not centrifuged but immediately aliquoted and stored at \(-80^\circ\text{C}\).
Each experimental situation consisted of 50 paired samples (a selected process versus the standard). There are five biomarkers that we consider here:

1. Urine Interleukin 18 (IL-18; LLD =12.5 pg/ml) contained 99 undetectable readings, yielding a 33% left-censoring rate

2. Neutrophil Gelatinase Associated Lipocalin (NGAL; LLD =4 ng/ml) contained 48 undetectable readings, yielding a 16% left-censoring rate

3. Kidney Injury Molecule-1 (KIM-1; LLD =59 pg/ml) contained 21 undetectable readings, yielding a 7% left-censoring rate

4. Liver Type Fatty Acid Binding Protein (L-FABP; LLD =3 ng/ml) contained 87 undetectable readings, yielding a 29% left-censoring rate

5. Cystatin C (LLD =0.005 mg/ml) contained 80 undetectable readings, yielding a 26.7% left-censoring rate

Tables 6.1-6.4 show the summary of the estimates and 95% confidence intervals of the CCC for three processes using the ML approach based on the bivariate lognormal model, the bivariate Weibull-Gamma model, the bivariate FGM-Weibull model and the bivariate Piecewise Uniform-Weibull model, respectively, and all are accounting for values below the LLD.

This data set consists of positive random variables and also it is right skewed. Figures 6.1-6.3 show the scatter plots of all five biomarkers between the measurements of the urinary biomarkers collected under a standard condition and values obtained in samples prepared under three different processes, labeled as Process A, Process B, and Process C, respectively.

Based on our simulation studies in Chapter 5, we expect that the ML method based on a new class of bivariate Weibull distributions will perform well if the two variables are not strongly correlated. Since the estimate of the CCC on the ML method based on the bivariate FGM-Weibull model works relatively well for most situations especially when sample size is greater than 50, the correlation is not too strong, and the percent censoring is between 25% to 60%, all of the results in Tables 6.1-6.4 suggest that only KIM-1 has the high level of agreement between the reference standard and all three different processes according to the bivariate lognormal and the bivariate FGM-Weibull, whereas the bivariate Weibull-Gamma and the bivariate
piecewise Uniform-Weibull display weak levels of agreement. In addition, L-FABP has the high level of agreement between the reference standard and Process B according to the bivariate lognormal and the bivariate FGM-Weibull, whereas the bivariate Weibull-Gamma and the bivariate piecewise Uniform-Weibull display weak levels of agreement. NGAL has the moderate level of agreement between the reference standard and processes A and B according to the bivariate lognormal and the bivariate FGM-Weibull, whereas the bivariate Weibull-Gamma and the bivariate piecewise Uniform-Weibull display weak levels of agreement. IL-18 and Cystatin C has the low level of agreement between the reference standard and all three different processes. In general, the bivariate lognormal can accommodate higher levels of agreement as seen in this example, whereas the bivariate Weibull models cannot
Figure 6.1. Scatter Plots of All biomarkers in Process A
Figure 6.2. Scatter Plots of All biomarkers in Process B
Figure 6.3. Scatter Plots of All biomarkers in Process C
### Table 6.1. Concordance Correlation Coefficients (and 95% Confidence Intervals) for the Three Processes Using the Bivariate Lognormal Model and Accounting for Values Below the LLD

<table>
<thead>
<tr>
<th>Processes</th>
<th>IL-18* (Initial 48 hours: 0°C vs -80°C)</th>
<th>NGAL* (Initial 48 hours: 25°C vs -80°C)</th>
<th>KIM-1* (Centrifuge vs No Centrifuge)</th>
<th>L-FABP*</th>
<th>Cystatin C*</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.83 (0.74, 0.92)</td>
<td>0.99 (0.98, 1.00)</td>
<td>0.99 (0.99, 1.00)</td>
<td>0.99 (0.98, 1.00)</td>
<td>0.97 (0.95, 0.99)</td>
</tr>
<tr>
<td>B</td>
<td>0.68 (0.51, 0.85)</td>
<td>0.98 (0.97, 0.99)</td>
<td>0.99 (0.99, 1.00)</td>
<td>0.96 (0.94, 0.98)</td>
<td>0.95 (0.91, 0.98)</td>
</tr>
<tr>
<td>C(Centrifuge vs No Centrifuge)</td>
<td>0.99 (0.98, 1.00)</td>
<td>0.99 (0.98, 1.00)</td>
<td>0.99 (0.99, 1.00)</td>
<td>0.99 (0.98, 1.00)</td>
<td>0.99 (0.99, 1.00)</td>
</tr>
</tbody>
</table>

*IL-18 (Interleukin 18) ; NGAL (Neutrophil Gelatinase Associated Lipocalin) ; KIM-1 (Kidney Injury Molecule-1) ; L-FABP (Liver Type Fatty Acid Binding Protein
### Table 6.2. Concordance Correlation Coefficients (and 95\% Confidence Intervals) for the Three Processes Using the Bivariate Weibull-Gamma Model and Accounting for Values Below the LLD

<table>
<thead>
<tr>
<th>Processes</th>
<th>IL-18*</th>
<th>NGAL*</th>
<th>KIM-1*</th>
<th>L-FABP*</th>
<th>Cystatin C*</th>
</tr>
</thead>
<tbody>
<tr>
<td>A (Initial 48 hours: 4 (^\circ)C vs -80 (^\circ)C)</td>
<td>0.24 (0.19, 0.29)</td>
<td>0.22 (0.20, 0.25)</td>
<td>0.25 (0.25, 0.26)</td>
<td>0.12 (0.08, 0.16)</td>
<td>0.14 (0.10, 0.19)</td>
</tr>
<tr>
<td>B (Initial 48 hours: 25 (^\circ)C vs -80 (^\circ)C)</td>
<td>0.21 (0.14, 0.28)</td>
<td>0.22 (0.19, 0.25)</td>
<td>0.25 (0.23, 0.27)</td>
<td>0.23 (0.18, 0.29)</td>
<td>0.15 (0.10, 0.20)</td>
</tr>
<tr>
<td>C(Centrifuge vs No Centrifuge)</td>
<td>0.17 (0.12, 0.21)</td>
<td>0.16 (0.12, 0.20)</td>
<td>0.25 (0.25, 0.26)</td>
<td>0.10 (0.05, 0.15)</td>
<td>0.12 (0.07, 0.17)</td>
</tr>
</tbody>
</table>

*IL-18 (Interleukin 18) ; NGAL (Neutrophil Gelatinase Associated Lipocalin ; KIM-1 (Kidney Injury Molecule-1 ; L-FABP (Liver Type Fatty Acid Binding Protein
Table 6.3. Concordance Correlation Coefficients (and 95% Confidence Intervals) for the Three Processes Using the Bivariate FGM-Weibull Model and Accounting for Values Below the LLD

<table>
<thead>
<tr>
<th>Processes</th>
<th>IL-18*</th>
<th>NGAL*</th>
<th>KIM-1*</th>
<th>L-FABP*</th>
<th>Cystatin C*</th>
</tr>
</thead>
<tbody>
<tr>
<td>A (Initial 48 hours: 4 °C vs -80 °C)</td>
<td>0.27</td>
<td>0.68</td>
<td>0.87</td>
<td>0.21</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>(0.18, 0.35)</td>
<td>(0.55, 0.81)</td>
<td>(0.76, 0.99)</td>
<td>(0.10, 0.32)</td>
<td>(0.14, 0.40)</td>
</tr>
<tr>
<td>B (Initial 48 hours: 25 °C vs -80 °C)</td>
<td>0.30</td>
<td>0.68</td>
<td>0.96</td>
<td>0.83</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>(0.16, 0.44)</td>
<td>(0.56, 0.80)</td>
<td>(0.83, 1.00)</td>
<td>(0.63, 1.00)</td>
<td>(0.19, 0.47)</td>
</tr>
<tr>
<td>C(Centrifuge vs No Centrifuge)</td>
<td>0.41</td>
<td>0.43</td>
<td>0.93</td>
<td>0.20</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>(0.26, 0.56)</td>
<td>(0.29, 0.58)</td>
<td>(0.82, 1.00)</td>
<td>(0.08, 0.32)</td>
<td>(0.07, 0.30)</td>
</tr>
</tbody>
</table>

*IL-18 (Interleukin 18) ; NGAL (Neutrophil Gelatinase Associated Lipocalin ; KIM-1 (Kidney Injury Molecule-1 ; L-FABP (Liver Type Fatty Acid Binding Protein
Table 6.4. Concordance Correlation Coefficients (and 95% Confidence Intervals) for the Three Processes Using the Bivariate Piecewise Uniform-Weibull Model and Accounting for Values Below the LLD

<table>
<thead>
<tr>
<th>Processes</th>
<th>IL-18*</th>
<th>NGAL*</th>
<th>KIM-1*</th>
<th>L-FABP*</th>
<th>Cystatin C*</th>
</tr>
</thead>
<tbody>
<tr>
<td>A (Initial 48 hours: 4 °C vs -80 °C)</td>
<td>0.19 (0.13, 0.24)</td>
<td>0.14 (0.10, 0.18)</td>
<td>0.21 (0.18, 0.24)</td>
<td>0.05 (0.02, 0.07)</td>
<td>0.08 (0.05, 0.12)</td>
</tr>
<tr>
<td>B (Initial 48 hours: 25 °C vs -80 °C)</td>
<td>0.14 (0.06, 0.22)</td>
<td>0.13 (0.09, 0.16)</td>
<td>0.21 (0.18, 0.25)</td>
<td>0.18 (0.12, 0.24)</td>
<td>0.08 (0.05, 0.11)</td>
</tr>
<tr>
<td>C(Centrifuge vs No Centrifuge)</td>
<td>0.06 (0.03, 0.09)</td>
<td>0.08 (0.05, 0.11)</td>
<td>0.21 (0.18, 0.25)</td>
<td>0.04 (0.02, 0.07)</td>
<td>0.06 (0.03, 0.08)</td>
</tr>
</tbody>
</table>

*IL-18 (Interleukin 18); NGAL (Neutrophil Gelatinase Associated Lipocalin); KIM-1 (Kidney Injury Molecule-1); LFABP (Liver Type Fatty Acid Binding Protein)
6.2 Asthma Clinical Data

The data for this example are taken from a crossover asthma clinical trial completed by the Asthma Clinical Research Network (ACRN). This data set was first analyzed by Deykin et al. [28]. The main objective of this trial was to determine whether the combination of the leukotriene receptor antagonist (LTRA) montelukast and the long-acting $\beta$-agonists (LABAs) could provide an effective therapeutic strategy for asthma. We are comparing this combination to the gold standard which is the combination of long-acting $\beta$-agonists (LABAs) and inhaled cortico-steroids (ICS). The primary outcome for this trial was time to treatment failure, which can be right-censored. Patients with physician-diagnosed asthma were randomized in a double-blinded fashion to treatment sequences in a $2 \times 2$ crossover design. The specific criteria that define the event of 'treatment failure' appear in Table 1 of Deykin et al. [28]. The treatments under investigation were a long-acting $\beta$-agonists (LABAs), an inhaled cortico-steroid (ICS), and a leukotriene receptor antagonist (LTRA), used in the following combinations within the crossover design as shown in Table 6.5. Frequency Counts and Cross Tabulation Table for ICS and LTRA are shown in Table 6.6.

Deykin et al. [28] investigated the analysis of a time-to-treatment failure outcome in a $2 \times 2$ crossover design via the method by France et al. [29], which reduces to McNemar’s test because the period effect was insignificant, and reported a p-value of 0.0008 for the treatment effect. Shvartsman et al. [30] also analyzed this data set via a proportional hazards regression model for analyzing a time-to-event outcome within the framework of an $s \times p$ crossover design, and summarized that the relative risk for treatment regimen LABA + ICS vs. LABA + LTRA is $\exp(-1.05) = 0.34$ with a 95% confidence interval of (0.09, 0.59), indicating that LABA + ICS significantly reduces the risk of treatment failure.

For this data set, the estimates and 95% confidence intervals of the difference of mean between LABA + LTRA and LABA + ICS for the ML approach based on the bivariate lognormal model and accounting for right-censored data equal to

$$\hat{\mu}_X - \hat{\mu}_Y = 435.73 \quad (95\% \ CI = (-1,325.35, 2, 196.82))$$

It shows that the time to treatment failure in LABA + LTRA is 436 days longer than from LABA + ICS, but notice that the confidence interval is very wide. There is not sufficient data in the ICS treatment regimen to get a good handle on this because of only 8 events in the
Table 6.5. The Two 14-week Randomized Treatment Periods

<table>
<thead>
<tr>
<th>Run-in Period 1 (4 weeks)</th>
<th>Treatment Period 1 (14 weeks)</th>
<th>Run-in Period 2 (4 weeks)</th>
<th>Treatment Period 2 (14 weeks)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence 1</td>
<td>Active LABA</td>
<td>Active LABA</td>
<td>Placebo ICS</td>
</tr>
<tr>
<td></td>
<td>Active ICS</td>
<td>Active ICS</td>
<td>Active LTRA</td>
</tr>
<tr>
<td></td>
<td>Placebo LTRA</td>
<td>Placebo LTRA</td>
<td>Placebo LTRA</td>
</tr>
<tr>
<td>Sequence 2</td>
<td>Active LABA</td>
<td>Active LABA</td>
<td>Active ICS</td>
</tr>
<tr>
<td></td>
<td>Placebo ICS</td>
<td>Active ICS</td>
<td>Active LTRA</td>
</tr>
<tr>
<td></td>
<td>Active LTRA</td>
<td>Placebo LTRA</td>
<td>Placebo LTRA</td>
</tr>
</tbody>
</table>

Long-acting β-agonists (LABAs), an inhaled corticosteroid (ICS), and a leukotriene receptor antagonist (LTRA)

Table 6.6. Frequency Counts and Cross Tabulation Table for ICS and LTRA

<table>
<thead>
<tr>
<th>Frequency</th>
<th>LTRA</th>
<th>ICS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>LTRA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>73</td>
<td>8</td>
</tr>
<tr>
<td>66.36%</td>
<td>7.27%</td>
<td>73.64%</td>
</tr>
<tr>
<td>1</td>
<td>25</td>
<td>4</td>
</tr>
<tr>
<td>22.73%</td>
<td>3.64%</td>
<td>26.36%</td>
</tr>
<tr>
<td>Total</td>
<td>98</td>
<td>12</td>
</tr>
<tr>
<td>89.09%</td>
<td>10.91%</td>
<td>100%</td>
</tr>
</tbody>
</table>

An inhaled corticosteroid (ICS), and a leukotriene receptor antagonist (LTRA)

ICS treatment regimen, see Table 6.6.
Chapter 7  
Future Studies

We have introduced the maximum likelihood method based on a new class of Weibull distributions to estimate the index of agreement, the concordance correlation coefficient, from left-censored data. First, we have explored available methods that are able to handle the left-censored data. Several techniques are presented for estimation of the index of agreement from data containing non-detectable values. The techniques investigated in Chapter 3 include methods of estimation with a left-censored value: (1) a maximum likelihood statistical method (2) deleting the pair data when either of them is below LLD (3) substitution of LLD for each non-detectable value (4) substitution of LLD/2 for each non-detectable value and (5) substitution of C × LLD for each non-detectable value, where LLD is the lower of limit of detection and C is the random number from the uniform[0, 1] distribution. The computer simulation studies in Chapter 3 show that the ML approach based on the bivariate lognormality assumption works best among all of the studied approaches. The advantages of the ML approach are that it is accurate (small relative bias) and accounts for the variability in the data set appropriately. Additionally, it uses all the available data for the statistical analysis. Therefore, in this work, we specifically focus on using the ML approach to estimate the index of agreement, CCC. In addition to the ML approach based on the bivariate lognormality assumption, we have considered different three bivariate distributions to apply to the maximum likelihood method: (1) Bivariate Weibull-Gamma distribution (2) Bivariate FGM-Weibull Distribution, and (3) Piecewise Uniform-Weibull distribution. The simulation results confirmed that overall in terms of accuracy, that is small relative bias, the estimator of CCC based on FGM-Weibull works relatively well in general cases when the correlation is not too strong even with the high percentage of censoring. For a skewed underlying
distribution with moderate or weaker correlation between two variables, the CCC estimated by a FGM-Weibull model is more robust. However, when the data are generated from the bivariate lognormal, the ML approach based on the bivariate lognormality assumption still performs best.

In this work, we have extended our bivariate density function to the situation with three random variables. Therefore, we can pursue extending these results for the Weibull-Gamma, the FGM-Weibull, and the Piecewise Uniform-Weibull distributions to the situation with four or more variables. In addition to the univariate Weibull distribution that we applied to our new class of survival functions, the univariate Gamma distribution is another positive-valued distribution that is widely used in many research areas that we can used to apply as well. This application can be apply to many situations. These topics for extension will be explored in future work. Another avenue for future work is to determine if there are modifications to these multivariate Weibull models that can accommodate stronger correlations.
Bibliography


Vita
Uthumporn Domthong

Education

• M.S. in Statistics, Western Michigan University, May 2009
• B.S. in Mathematics, Khon Kean University, Thailand March 2005

Academic Experience at The Pennsylvania State University

• Spring 2010 - Summer 2012 Teaching Assistant at the Statistics Department
• Spring 2013 - Teaching Assistant at the Public Health Science Department
• Summer 2013 - Fall 2014 Research Assistant at the Public Health Science Department